UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Matjaž Gulja

Analiza računalniškega omrežja na primeru Dijaškega doma Tabor

DIPLOMSKO DELO NA

VISOKOŠOLSKEM STROKOVNEM ŠTUDIJU

Mentor: doc. dr. Miha Moškon

Ljubljana, 2014

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika naloge: *Analiza računalniškega omrežja na primeru Dijaškega doma Tabor*

Kandidat naj v diplomski nalogi izbere metodologijo in preuči orodja za izvajanje analize računalniških omrežij. Analizo naj v praksi izvede na zgledu računalniškega omrežja Dijaškega doma Tabor. Na podlagi opravljenega dela naj kandidat predlaga možnosti za izboljšanje delovanja obravnavanega omrežja in poda smernice za razvoj omrežja v prihodnosti.

I Z J A V A O A V T O R S T V U

DIPLOMSKEGA DELA

Spodaj podpisani **Matjaž Gulja**, z vpisno številko **63010350**,

sem avtor diplomskega dela z naslovom:

Analiza računalniškega omrežja na primeru Dijaškega doma Tabor

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom **doc. dr. Miha Moškona,**
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela,
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki »Dela FRI«.

V Ljubljani, dne 19. 9. 2014 Podpis avtorja: _____

ZAHVALA

Za pomoč in usmerjanje pri izdelavi diplomske naloge se zahvaljujem mentorju doc. dr. Mihi Moškonu. Obenem se za nesebično pomoč in podporo v času pisanja naloge zahvaljujem tudi prijateljem Damjanu, Urošu in Denisu.

Zahvalil bi se še staršem in punci, ki so mi ob strani stali ves čas študija.

Hvala!

KAZALO

POVZETEK

V nalogi smo preučili različna orodja za analiziranje prometa v računalniških omrežij in jih uporabili pri analizi omrežja Dijaškega doma Tabor. Meritve smo opravljali v času turistične sezone in na začetku šolskega leta, ko turiste zamenjajo dijaki. Primerjali smo navade obeh populacij ter iskali skupne vzorce. Med drugim smo opravili analizo prenosa datotek preko medcelinskih povezav in ugotavljali vpliv nastavitev protokola TCP na samo hitrost prenosa. Za analize smo v VMware ESXi uporabljali programski usmerjevalnik pfSense ter razvili dodatne Python skripte, ki razčlenjujejo dnevniške datoteke za nadaljnjo obdelavo z SQL poizvedbami v PostgreSQL orodju in z analitičnim orodjem MicroStrategy Analytics.

Ključne besede: analiza omrežja, analiza prometa, pfSense, VMware, ESXi, Python, PostgreSQL, MicroStrategy Analytics

ABSTRACT

Thesis presents the overview of various computer traffic analysis tools. The case study of the analysis is performed on the Tabor dormitories' computer network. Two different populations were observed, i.e. the residents during the tourist season and during the school year. Common patterns between the groups were analysed. The analysis of file transfers via intercontinental connections was conducted additionally to observe the influence of TCP settings on the download speed. The pfSense and VMware ESXi were used for the environment configuration. Log files were parsed with custom Python scripts and processed with SQL queries in PostgreSQL and with MicroStrategy Analytics.

Keywords: network analysis, traffic analysis, pfSense, VMware ESXi, Python, PostgreSQL, MicroStrategy Analytics

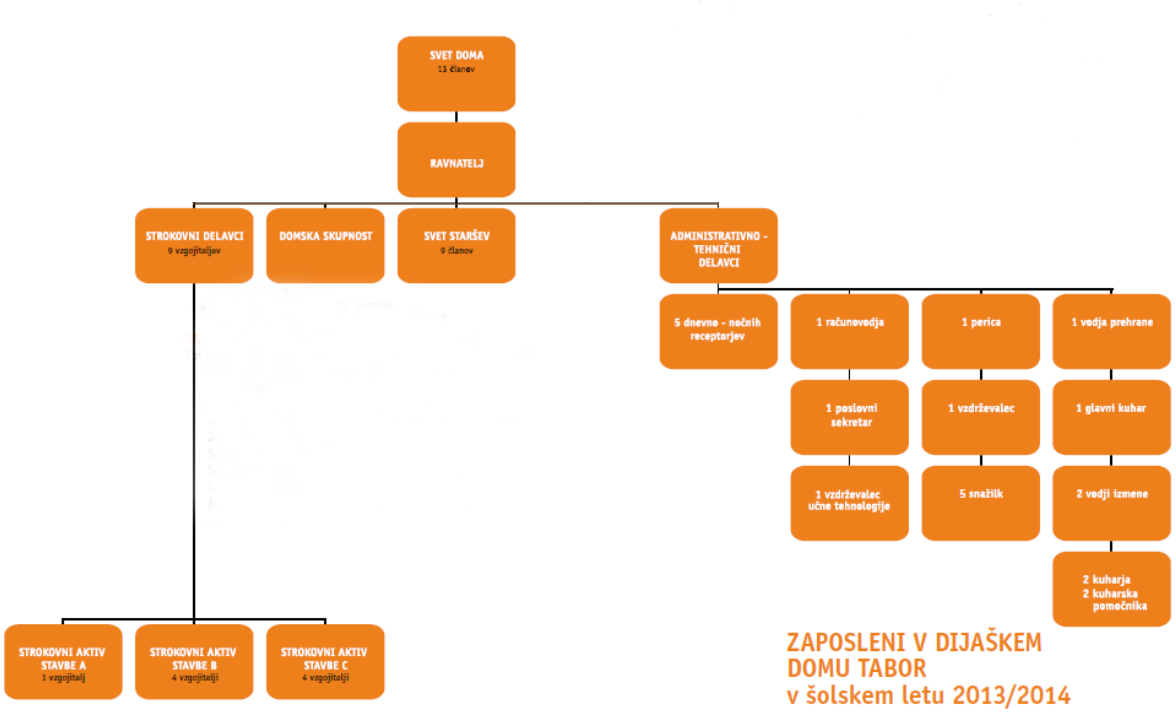
1 UVOD

Dijaški domovi imajo v slovenskem šolskem prostoru večletno zgodovino in kot taki tudi predstavljajo nepogrešljiv del šolskega sistema. Čeprav je njihov namen predvsem zadostitev potreb dijakov po hrani, namestitvi in zdravstveni zaščiti, pa so le-ti po svoji osnovi prvenstveno vzgojne ustanove, kot taki pa del vzgojno-izobraževalnega sistema [1].

Informacijski sistem predstavlja pomemben del vzgojno-izobraževalnega procesa, saj dijakom nudi prostor, kjer se izobražujejo, družijo, ustvarjajo in razmišljajo. Zelo pomembno je torej, da dijaški dom kot del vzgojno-izobraževalnega sistema v današnjem času svojim stanovalcem nudi tudi zadostno informacijsko podporo.

Dijaški dom Tabor je vzgojno-izobraževalna ustanova, ki je bila ustanovljena s strani države. Njegovi začetki segajo že v leto 1948, ko se je število dijakov z željo po tehniški izobrazbi skokovito povečalo, s tem pa se je dvignila tudi potreba po tovrstni ustanovi. Tekom zgodovine se je dom nekajkrat preimenoval, pod zdajšnjim imenom pa ga poznamo zadnjih 40 let.

Dijaški dom Tabor je sestavljen iz treh stavb (A, B in C). Zaposleni ter dijaki v njem so porazdeljeni po vseh stavbah, zato je samo omrežje precej razvejano. Organizacijsko se delavci delijo na vodstveni, strokovni, administrativni ter tehnični sektor (*slika 1*). Vodstveni oddelek predstavlja ravnatelj, ki je skupaj s celotnim administrativnim sektorjem in tremi strokovnimi sodelavci lociran v stavbi A. Strokovni kader predstavlja tretjino vseh zaposlenih, ki se delijo na vzgojitelje ter svetovalno službo. Svoje prostore imajo v vseh stavbah, strokovni kader je obenem tudi največji uporabnik omrežne infrastrukture, najmanj pa jo uporabljajo tehnični delavci (čistilke, hišnik, kuhinjsko osebje). Seveda pa dijaški dom ne bi obstajal brez dijakov. V zadnjih 15 letih se je število dijakov več kot razpolovilo, iz 600 na približno 250, nekaj na račun višjih standardov, nekaj pa tudi zaradi zmanjšanja šolske populacije.



Slika 1: Organizacijska struktura zaposlenih

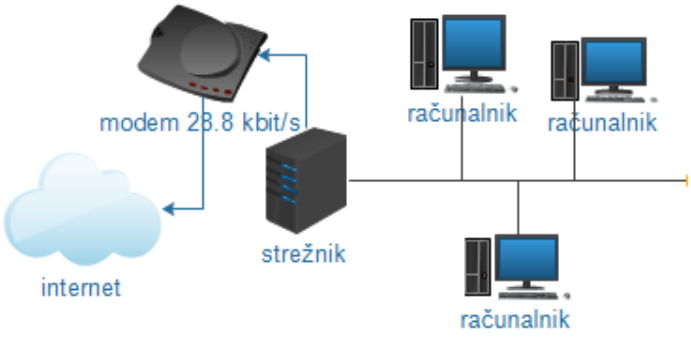
V tem delu predstavimo informacijski sistem Dijaškega doma Tabor in izvedemo analizo obremenjenosti omrežja ter analizo navad uporabnikov, predvsem s stališča računalniškega omrežja, ki ga informacijski sistem nudi. Na podlagi rezultatov ocenimo primernost omrežne infrastrukture in podamo smernice za potencialen razvoj omrežja v prihodnosti.

V poglavju 2 opišemo omrežno infrastrukturo Dijaškega doma Tabor. Poglavje 3 predstavlja metodologijo opravljenih meritev, poglavje 4 pa predstavitev rezultatov obdelave dnevniških zapisov iz posredniškega (*angl. proxy*) strežnika, požarnega zida in meritev hitrosti mednarodnih prenosov.

2 OMREŽNA INFRASTRUKTURA

2.1 Kratka zgodovina omrežne infrastrukture

Prvi korak k računalniškem opismenjevanju v domu sega v daljno leto 1988 z nakupom prvega računalnika za potrebe računovodskih del. Leta 1991 smo od Fakultete za matematiko in fiziko dobili 11 računalnikov Iskra Partner s priloženim operacijskim sistemom in nekaj dodatne programske opreme, med katero sta bila zanimiva predvsem urejevalnik besedil Wordstar in programski jezik Pascal, ki so ga uporabljali dijaki srednje šole za računalništvo. Pozneje smo s pomočjo sponzorja (Kovinoplastika Lož) in Ministrstva za šolstvo in šport pridobili nove računalnike (procesor 386 in 486), ki so jih dijaki postopoma vse bolj uporabljali za šolsko delo. Potreba po računalnikih je naraščala iz leta v leto, saj so bili v domu pretežno dijaki srednjih tehniških šol. Dejansko pa je računalništvo v dijaškem domu popolnoma zaživelo v letu 1995 z izgradnjo nove, večje računalnice, ki je služila tudi kot centralna lokacija za računalniško opismenjevanje vzgojiteljev vseh dijaških domov Slovenije. Takratno omrežje je bilo sestavljeno iz 12 računalnikov, v omrežje so bili med seboj povezani s koaksialnim kablom. To pomeni, da je bil vsak izmed računalnikov zaporedno povezan preko Bayonet Neill-Concelman (krajše *BNC*) veznega elementa, na koncu verige pa je bil obvezen zaključni (*BNC-T*) člen. Povezava z internetom se je vzpostavljala s pomočjo 28.8 kbit/s modema s klicno povezavo do ARNES omrežja (*slika 2*).



Slika 2: Povezovanje z modemom in BNC veznimi členi

2.2 Omrežna infrastruktura danes

Potrebe po vse hitrejšem dostopu do interneta so eksponentno naraščale, zato smo leta 1998 opustili tehnologijo modemske povezave, ter se preko najetega voda Telekoma Slovenije s pasovno širino 256 Kb/s priključili v omrežje ARNES (*Slovenska akademska in raziskovalna mreža Slovenije*) in s tem dobili neprekinjen dostop do interneta. Takšna hitrost povezave je že omogočala kvalitetno izvajanje raznih računalniških tečajev za dijake in zaposlene.

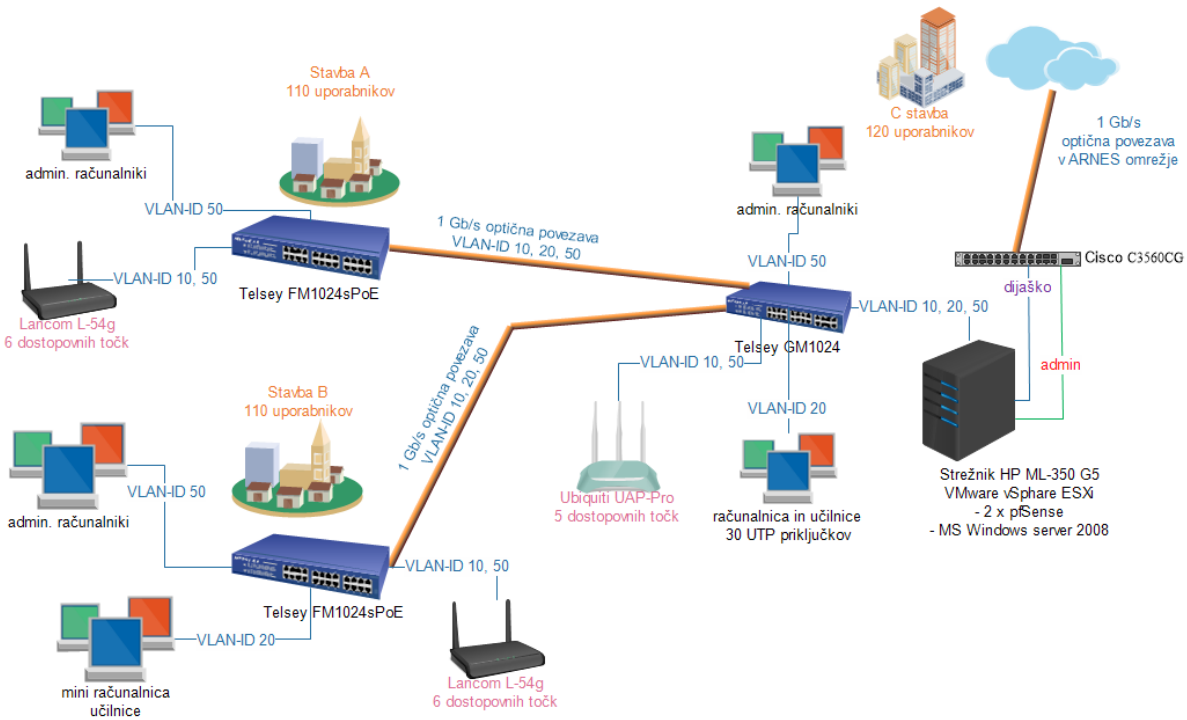
Ob zamenjavi povezave smo razširili in posodobili tudi notranjo omrežno infrastrukturo. V glavni stavbi C smo vzpostavili računalniški center ter ga s koaksialnim kablom povezali z računalnico in drugima stavbama. Računalniki v vseh prostorih doma so bili v omrežje postavljeni s pomočjo koncentratorja (*angl. hub)*. Kmalu je sledila naslednja nadgradnja, saj je 256 Kb/s povezava postala ozko grlo. S pomočjo Telekoma Slovenije in ARNES-a smo pridobili svojo najeto optično povezavo iz Telekomovega vozlišča na Cigaletovi ulici, priključeno v ARNES omrežje. Prvo leto je bila povezava s Telekomovim vozliščem 10 Mb/s, naslednje leto pa že 100 Mb/s. Za pretvorbo optičnega v električni signal smo uporabljali Telekomov media pretvornik. Tako je bilo vse do leta 2013, ko smo zaprosili in tudi dobili nadgradnjo na 1 Gb/s povezavo, kjer ni več potrebe po pretvorniku, saj je optično vlakno pripeljano neposredno v Cisco stikalo (*slika 3*).

Slika 3: Povezava v ARNES omrežje z VLAN-i

ARNES nam je dodelil dve ločeni omrežji. Za potrebe dijakov v učilnicah, računalnici in brezžičnem omrežju so nam dodelili blok za priključitev 126 naprav, medtem ko vsi zaposleni (pisarne, vzgojitelji, tehnično in administrativno osebje) uporabljamo ločeno omrežje z možnostjo priključitve 30 naprav. Ker pride iz ARNES-ovega stikala v naš računalniški center

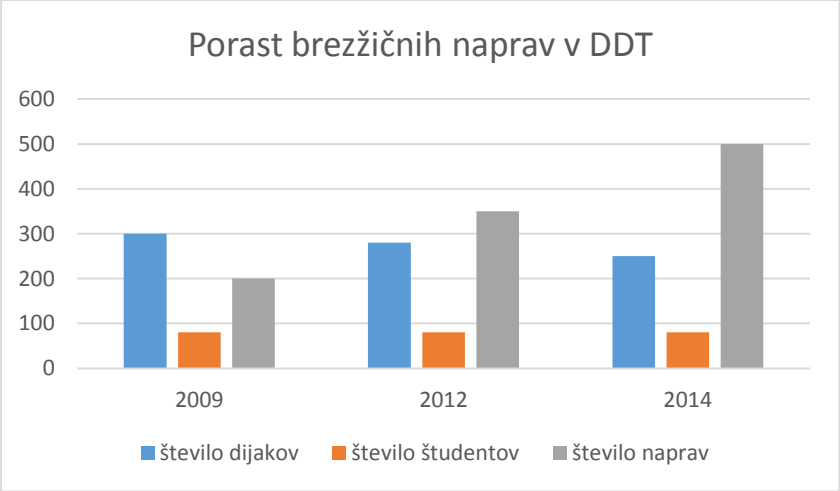
samo en optični kabel, se omrežji ločita s pomočjo navideznih omrežij (*angl. VLAN*) ali tako imenovanega standarda IEEE 802.1Q [2]. Posamezen vhodni ali izhodni podatkovni paket natančno ve, po katerem navideznem omrežju mora potovati, saj ima vsako omrežje svojo unikatno oznako (*angl. VLAN-ID*). Isti standard uporabljamo tudi v notranjem omrežju. Cilj tega je, da dijaki ne morejo neposredno dostopati do podatkov v administrativnem omrežju.

V času prehoda na 10 Mb/s optično povezavo, smo preuredili tudi lokalno omrežje *(angl. LAN)*. Povezave s koaksialnim kablom med stavbami in računalniškim centrom smo zamenjali z optičnim vlaknom. Stare koncentratorje v posameznih vozliščih v stavbah so zamenjala modernejša 100 Mb/s stikala Telsey FM1024SPoE s 24 100 Mb/s priključki (*angl. ports*) in s štirimi gigabitnimi priključki. En priključek s SFP (*angl. Small Form-factor Pluggable*) modulom služi kot medpovezava z računalniškim centrom, v katerem je locirano stikalo Telsey TS-GM1024 s 24 gigabitnimi priključki. Od teh se lahko štirje uporabijo v povezavi s SFP modulom. Trenutno uporabljamo le dva SFP porta, za optično povezavo do obeh stavb. Optično vlakno, ki povezuje obe stavbi je večrodovno (*angl. multimode*), kar pomeni, da lahko po njem potuje več valovnih dolžin od 900 do 1000 nm, namenjen pa je povezovanju krajših razdalj (*slika 4*). [3]



Slika 4: Omrežje v DD Tabor

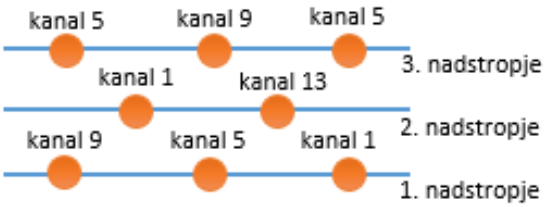
Z nakupom novih dostopovnih točk smo leta 2009 v vseh stavbah doma vzpostavili tudi brezžično omrežje. Do takrat smo namreč možnost brezžičnega povezovanja imeli le v skupnih prostorih, kar je dijake pri pridobivanju novih znanj zelo omejevalo. Že ob nakupu Lancom L-54g dostopnih točk smo se zavedali, da bodo te za strmo rast mobilnih naprav in čedalje večje količine prenesenih podatkov hitro postale premalo zmogljive. V tistem letu se je namreč že oblikoval nov standard 802.11n, ki omogoča višje hitrosti prenosa, vendar ta s strani organizacije IEEE (*angl. Institute of Electrical and Electronics Engineers*) še ni bil dokončno potrjen. Ob zagonu novega omrežja je bilo v dijaškem domu vpisanih 300 dijakov in 80 študentov, v sistem pa je bilo registriranih zgolj približno 200 brezžičnih naprav, večinoma prenosnikov. V šolskem letu 2013/2014 pa smo imeli registriranih že 500 naprav in to kljub dejstvu, da je število dijakov padlo na 250, število študentov pa je ostalo na približno isti ravni, torej 80 (*slika 5*).



Slika 5: Prikaz porasta brezžičnih naprav

Ob vse bolj intenzivni uporabi omrežja smo prišli do spoznanja, da je, predvsem v skrajnih delih stavbe oz. sobah, prisotna težava z zagotavljanjem pokritosti signala in s tem nemotenega dostopa do spletnih vsebin. Odločili smo se, da bomo po stavbah postopoma nakupili zmogljivejše točke proizvajalca Ubiquiti Networks s karakteristikami, ki zadostujejo trenutnim potrebam. Dosedanje točke iz ene stavbe bomo nato dodali v druge stavbe in s tem izboljšali pokritost. Do sedaj smo jih na treh nadstropjih imeli postavljene po sistemu 2, 1, 2, po novem pa bodo dodane še 3 točke in razporejene na način 3, 2, 3 (*slika 6*). Kljub večji gostoti točk tako ne bo prihajalo do prekrivanja signala posameznih kanalov in s tem

povezanih motenj pri prenosu, saj so dostopovne točke nastavljene na točno določene kanale, ki se med seboj ne prekrivajo (1, 5, 9 in 13 – *slika 6*).



Slika 6: Način postavitve točk in nastavitev kanalov

Vsaka dostopovna točka z navideznimi omrežji razdeli omrežje na dijaško in administrativno. Omrežje za potrebe zaposlenih je zaščiteno s standardom IEEE 802.11i (WPA 2 s TKIP/AES kodiranjem). Dijaško pa je dostopno brez gesla (*angl. open network*), za pridobitev veljavnega IP naslova in s tem dostopa do interneta je potreben vnos fizičnega naslova mrežne kartice (*angl. Media Access Control,* krajše *MAC*) v usmerjevalnik, ki ga poganja sistem pfSense. Vsak dijak s svojim imenom, številko sobe in MAC naslovom omrežne kartice registrira svoje naprave v sistem. S tem imamo olajšan nadzor nad nedovoljenim početjem dijakov. V času turistične dejavnosti (poletne počitnice) tak način identifikacije izklopimo in našim gostom omogočimo prosto uporabo omrežja.

2.2.1 Stavba A

V komunikacijski omari te stavbe se nahaja stikalo Telsey FM1024SPoE. Stikalo je z navideznimi omrežji razdeljeno na 3 podomrežja. VLAN-ID 10 je namenjen dijaškemu brezžičnemu dostopu, VLAN-ID 20 povezuje dijaške fizične UTP (*angl. Unshielded Twisted Pair*) priključke, VLAN-ID 50 pa je namenjen napravam na delovnih mestih in brezžičnemu dostopu za zaposlene. V tej stavbi imamo štiri učilnice s po dvema UTP priključkoma ter mini računalnico s šestimi priključki. Zaposlenim v tej stavbi nudimo šest delovnih postaj v petih pisarnah, tri mrežne tiskalnike ter skener. Za pokritost brezžičnega signala v stavbi skrbi šest dostopovnih točk, za njihovo napajanje z elektriko pa skrbi stikalo Telsey preko standarda IEEE 802.3af (*angl. Power over Ethernet,* krajše *POE)*. Starejše dostopovne točke Lancom delujejo samo na frekvenci 2.4 GHz, novejše Ubiquiti pa sočasno tudi na frekvenci 5 GHz.

2.2.2 Stavba B

V komunikacijski omari te stavbe se nahaja stikalo Telsey FM1024SPoE. Razdeljeno je na 3 navidezna podomrežja z VLAN-ID 10, 20 in 50. Delavcem stikalo služi za dostop do omrežja na sedmih delovnih mestih v štirih pisarnah ter za priklop dveh mrežnih tiskalnikov, dijakom pa uporabo omrežja omogoča na petih računalnikih v dijaški mini računalnici, šestih priključkih v učilnicah in pet dostopovnih točkah. Pred uvedbo brezžičnega omrežja smo v nekaj sob poskusno napeljali UTP kabel za dostop do interneta, vendar smo projekt zaradi prednostne uvedbe brezžičnega sistema nato zaustavili.

2.2.3 Stavba C

V tej stavbi se nahaja računalniški center z glavno komunikacijsko omaro. Iz Telekomovega centra na Cigaletovi ulici, kjer se nahaja tudi ARNES-ovo vozlišče, je v omaro pripeljan najeti par neosvetljenih optičnih vlaken (*angl. dark fiber*). Kot šolska organizacija smo od ARNES-a v uporabo dobili gigabitno stikalo Cisco C3560CG-8TC-S v katerega je preko SFP modula *GLC-LH-SM* priključeno *singlemode* optično vlakno. Stikalo, ki ga upravlja ARNES, je s pomočjo VLAN-ov razdeljeno na administrativno in dijaško podomrežje. Obe omrežji sta povezani s strežnikom HP ML350-G5, na katerem je nameščen VMware vSphere ESXi, v njem pa programski usmerjevalnik pfSense. PfSense je z VLAN-ID 10, 20 in 50 povezan s stikalom Telsey TS-GM1024. Slednje služi kot glavno stikalo, preko katerega poteka povezava z delovnimi mesti v C stavbi, glavno računalnico, učilnicami in video nadzornim sistemom. Stikalo skrbi tudi za optično povezavo z drugimi stavbami. V tej stavbi imamo pet delovnih mest, video nadzorni sistem, deset UTP priključkov v učilnicah, tri mrežne tiskalnike, 16 računalnikov za dijaške potrebe v glavni računalnici ter pet brezžičnih Ubiquiti UAP-Pro dostopovnih točk.

2.3 Strojna oprema

Strojna oprema v dijaškem domu je bila večinoma pridobljena v sklopu raznih sofinanciranj s strani Ministrstva za šolstvo, saj sami tolikšnega finančnega bremena ne bi zmogli. V letu 2008 smo se na njihov zadnji uspešen razpis za sofinanciranje omrežne opreme prijavili in tako dobili del sredstev za nakup in posodobitev mrežne tehnologije. Nabavili smo tri nova stikala znamke Telsey, nov strežnik Hewlett Packard serije ProLiant, v celoti smo prenovili brezžično infrastrukturo z nabavo 15 dostopovnih točk L-54g podjetja Lancom Systems. Za

upravljanje dostopovnih točk skrbi Lancom WLC-4025 nadzorna naprava. Te dostopovne točke sedaj počasi nadomeščamo z novejšimi podjetja Ubiquiti, natančneje z modelom UAP-Pro. Ob prehodu na 1 Gb/s optično povezavo smo od ARNES-a prevzeli tudi stikalo Cisco WS-C3560CG-8TC-S [4]. Za brezprekinitveno napajanje naprav v računalniškem centru skrbi Socomec Modulys 3kVA UPS, ki zagotavlja vsaj 20 minutno avtonomijo. Osebne računalnike smo v zadnjih letih posodabljali, tako da imamo sedaj vse z vsaj dvojedrnim procesorjem in 2 GB spomina. Za varnost podatkov skrbi NAS naprava Qnap TS-879 Pro s petimi diski povezanimi v polje RAID 5.

2.3.1 Stikala v komunikacijskih omaricah po stavbah

Cisco WS-C3560CG-8TC-S [4]

Glavno stikalo v dijaškem domu, ki ga upravlja ARNES, skrbi za nemoteno povezavo v njihovo omrežje. Njegova osnovna funkcija je, da s podporo standardu IEEE 802.1Q omrežje razdeli na dva dela: administrativno in dijaško. Stikalo nam s pomočjo SFP *singlemode* optičnega modula zagotavlja 1 Gb/s povezavo z internetom.

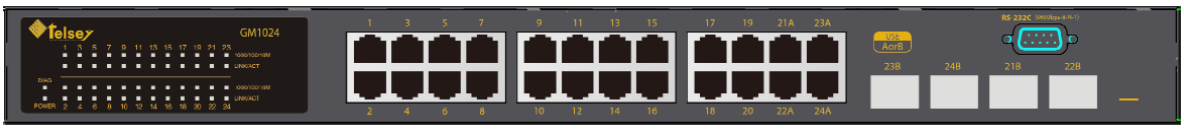
Telsey FM1024sPoE (*slika 7*) [5]

Stikalo deluje na povezovalni plasti OSI referenčnega modela (*angl. layer 2 - data*) [6]. Omogoča priklop v omrežje do 24 naprav s hitrostjo 10/100 Mb/s, ima 2 priključka za SFP modul (eden je uporabljen za optično povezavo z računalniškim centrom) in 2 priključka za naprave z 10/100/1000 Mb/s hitrostjo. Vseh 24 priključkov omogoča napajanje drugih naprav preko omrežnega kabla - po standardu IEEE 802.3af (*angl. Power over Ethernet,* krajše *PoE*), ki zagotavlja moč do 15.4 Watt [*7*]. Stikalo dovoljuje tudi združevanje vrat za hitrejšo povezavo (*link aggregation*) in razdelitev omrežja s pomočjo IEEE 802.1Q standarda (*VLAN*). Za naprednejšo kontrolo dostopa do omrežja lahko omogočimo dostop na podlagi posameznih vrat preko standarda IEEE 802.1X (*angl. Port-based Network Access Control,* krajše *PNAC*) in RADIUS (*angl. Remote Authentication Dial In User Service*) prijavo. RADIUS nam namreč omogoča centralizirano avtentikacijo, avtorizacijo in upravljanje z računom. Stikalo lahko upravljamo preko spletnega ali ukaznega vrstičnega vmesnika (*angl. command line interface,* krajše *CLI*) preko zaščitene SSH povezave.

Slika 7: Stikalo Telsey FM1024sPoE/2XG

Telsey TS-GM1024 (*slika 8*) [8]

To *layer 2* stikalo ima 24 priključkov, ki lahko delujejo s hitrostjo 10/100/1000 Mb/s. Ima možnost priklopa štirih SFP modulov, vendar lahko naenkrat deluje le SFP ali eden od pripadajočih štirih priključkov. *Layer 2* nam omogoča uporabo VLAN-ov, združevanje vrat za hitrejšo povezavo (*angl. link aggregation*), podpira *Jumbo Frames* do 9216 bajtov, uporabo IEEE 802.1D/s/W *spanning tree* protokola. Omogoča tudi kontrolo dostopa po standardu IEEE 802.1X in RADIUS avtentikacijo, na posameznem portu pa lahko na štirih nivojih tudi prioritiziramo promet po standardu IEEE 802.1p. Upravljamo ga preko spletnega ali ukaznega vrstičnega vmesnika (*CLI*).



Slika 8: Telsey GM1024 stikalo

2.3.2 HP Proliant ML-350 G5 strežnik

Omenjeni strežnik je za naš dijaški dom velika pridobitev, saj smo do takrat kot strežnik uporabljali navaden računalnik, na katerega je bil nameščen MS Windows Server 2003. Za namenski strežnik smo se posledično odločili, ker nam ta s svojo arhitekturo omogoča uporabo virtualnega okolja. Njegove glavne karakteristike so [9]:

- osnovna plošča z dvema Intel Xeon E5420 2.50 GHz procesorjema,
- 16 GB PC2-5300 DDR2 spomina,
- 8 SAS 2.5" diskov s hitrostjo vrtenja 10.000 obratov/min (4 x 146 GB v RAID 5 polju in 4 x 300 GB v RAID 5 polju),
- RAID (*angl. Redundant Array of Inexpensive Disks*) krmilnik E200i/128 BBWC z dodano baterijo za obvarovanje podatkov v primeru izpada električne energije,

- vgrajena gigabitna mrežna kartica z 1 vrati ter dodatna gigabitna HP mrežna kartica s 4 vrati,
- dodaten napajalnik za neprekinjeno delovanje v primeru okvare enega od njiju.

2.3.3 Lancom dostopne točke in nadzorna naprava

Podjetje Lancom System iz Nemčije že več kot deset let izdeluje profesionalno omrežno opremo, na kar nas opominjajo tudi številne prejete nagrade ter več kot milijon nameščenih naprav. S pomočjo sofinanciranja smo kupili 15 dostopnih točk L-54g (*slika 9*). Te točke je priporočal tudi ARNES, saj so jih testirali in so pripravljene za morebitno nadgradnjo na izobraževalno brezžično omrežje Eduroam. Njihove ključne lastnosti so [10]:

- delujejo v 2.4 GHz frekvenčnem območju,
- podpirajo standard IEEE 802.11b/g, s hitrostjo 54 Mb/s,
- ločevanje uporabnikov s pomočjo VLAN (IEEE 802.1q) ali več SSID oznak omrežja,
- napajanje preko podatkovnega kabla – PoE (IEEE 802.3af),
- 2 zamenljivi zunanji anteni,
- varnost prenosov po standardih IEEE 802.11i (WPA2-Personal), IEEE 802.1x/EAP, LEPS (WPA2-Enterprise) ter AES kodiranje,
- zagotavljanje kakovosti storitve (*Quality of service*, krajše *QoS*),
- prehajanje med sosednjimi točkami brez prekinitve prenosa (*roaming*).

Zaradi lažjega upravljanja takšnega števila dostopnih točk, smo kupili tudi nadzorno napravo WLC-4025. Preko enostavnega vmesnika za upravljanje omogoča nadzor do 25 točk. Vse spremembe nastavitev v brezžičnem omrežju avtomatsko razpošlje vsem točkam.

Slika 9: Dostopna točka Lancom L-54g

2.3.4 Ubiquiti dostopne točke UAP-Pro

Pri posodobitvi brezžičnega omrežja smo se odločili za znamko Ubiquiti, in sicer za njihove naprave iz razreda UniFi. Slednji predstavlja segment brezžičnih naprav, ki je osredotočen na zahtevnejša poslovna okolja (*enterprise wifi*). Omogoča neomejeno razširljivost, tudi z novejšimi napravami, saj so na trg že lansirali naprave z IEEE 802.11ac standardom. Vse dostopne točke se upravlja v brezplačnem programskem upravljalniku (*angl. software controller*). Program omogoča enostavno upravljanje in pregleden nadzor nad dostopovnimi točkami, saj lahko vstavimo svojo skico stavbe in na njo postavimo točke. Z njim lahko vodimo tudi statistiko uporabe omrežja (število uporabnikov na posamezni napravi, količina prenosa podatkov, moč signala, itd). Prav tako pa nam omogoča prehajanje med sosednjimi točkami brez prekinitve povezave (*angl. roaming*).

Naše brezžično omrežje bomo nadgrajevali s točkami UAP-Pro (*slika 10*). Deluje na dveh frekvenčnih pasovih istočasno. Na 2.4 GHz pasu podpira tri sočasne neodvisne podatkovne toke, s pomočjo treh vgrajenih anten, ki podpirajo 3x3 Multiple Input Multiple Output (MIMO). Hitrosti prenosa na tej frekvenci lahko dosežejo 450 Mb/s. 5 GHz frekvenčno delovanje pa z dvema vgrajenima antenama omogoča dva sočasna podatkovna toka (2x2 MIMO) ter hitrost do 300 Mb/s. V idealnih pogojih (brez ovir) lahko signal doseže 122 metrov. Omrežni priključek omogoča 10/100/1000 Mb/s prenose ter napajanje preko PoE standarda, ki je tudi edini možni način napajanja. Naprava podpira standarde IEEE 802.11a/b/g/n, od tega lahko 802.11a/n deluje tudi v 5 GHz frekvenčnem pasu. Za posamezni frekvenčni pas lahko kreiramo več omrežij s pomočjo VLAN-ov ali več SSID oznak omrežja. Omrežje lahko zaščitimo v vsemi popularnimi standardi (WEP, WPA-PSK, WPA-TKIP, WPA2 AES, 802.11i), omogoča pa tudi zagotavljanje kakovosti storitve (QoS), vendar le na nivoju omejitve prenosa do posameznega uporabnika. [11]

Slika 10: Dostopna točka Ubiquiti UAP-Pro

2.4 Programska oprema

V dijaškem domu stremimo k temu, da uporabljamo čim več odprtokodne programske opreme. Vse licenčne programe, ki jih uporabljamo smo pridobili s pomočjo sofinanciranja Ministrstva za šolstvo, posebnih popustov za vzgojno-izobraževalne zavode ali najema. Ker so se v času ekonomske krize razpisi s strani ministrstva praktično ustavili, moramo ob samostojnem nakupu osebnega računalnika kupiti tudi najosnovnejšo verzijo Microsoft Windows operacijskega sistema. Operacijski sistem pa lahko potem brezplačno nadgradimo na Enterprise različico. Ministrstvo je marca 2014 s podjetjem Microsoft podpisalo novo pogodbo Microsoft Enrollment for Education Solutions (pogodba EES). Na podlagi te pogodbe smo upravičeni do letnega najema njihovih dodatnih produktov, pri nas imamo tako najet Microsoft Windows Server Standard 2008, za katerega plačujemo simboličnih 55 EUR.

Že v letu 2008 ob nakupu strežnika HP ML-350 G5 smo razmišljali v smeri virtualizacije operacijskih sistemov, nižanja stroškov električne energije in nakupa več namenskih računalnikov ter poenostavitve upravljanja. V ta namen uporabljamo program VMware vSphere Hypervisor, poznan tudi kot VMware ESXi, ki je na uradni spletni strani podjetja VMware na voljo brezplačno. V tem virtualnem okolju imamo nameščenih več sistemov, med katerimi so najbolj izpostavljeni pfSense, brezplačni zmogljiv programski usmerjevalnik, ter Microsoft Windows Server 2008, strežnik namenjen zaposlenim.

2.4.1 VMware vSphere Hypervisor

Virtualizacija je tehnologija, ki svoj pomen na IT področju povečuje iz dneva v dan. Omogoča številne prednosti, kot so večji izkoristek strojne opreme, nižanje obratovalnih stroškov, večjo varnost in zanesljivost podatkov, poenostavljeno in centralno upravljanje sistemov, enostavnejšo nadgradnjo ali zamenjavo komponent. Poglavitna lastnost virtualizacije pa je, da lahko na en fizičen računalnik namestimo več različnih operacijskih sistemov, ki tečejo vsak v svojem virtualnem stroju. Ti virtualni stroji omogočajo, da si vsi operacijski sistemi delijo skupno strojno opremo, delujejo sočasno, neodvisno drug od drugega. Tovrstno delovanje omogoča nadzorni sistem hipervisor (*angl. hypervisor*), ki ga lahko namestimo neposredno na strojno opremo (*bare metal hypervisor, slika 11*) ali pa ga namestimo znotraj že nameščenega operacijskega sistema. Zelo razširjeni programi za takšen način virtualizacije so VMware Workstation (oz. zastonjska okleščena verzija VMware Player), Oracle VirtualBox in KVM (Kernel-based Virtual Machine) za Linux/Unix sisteme.

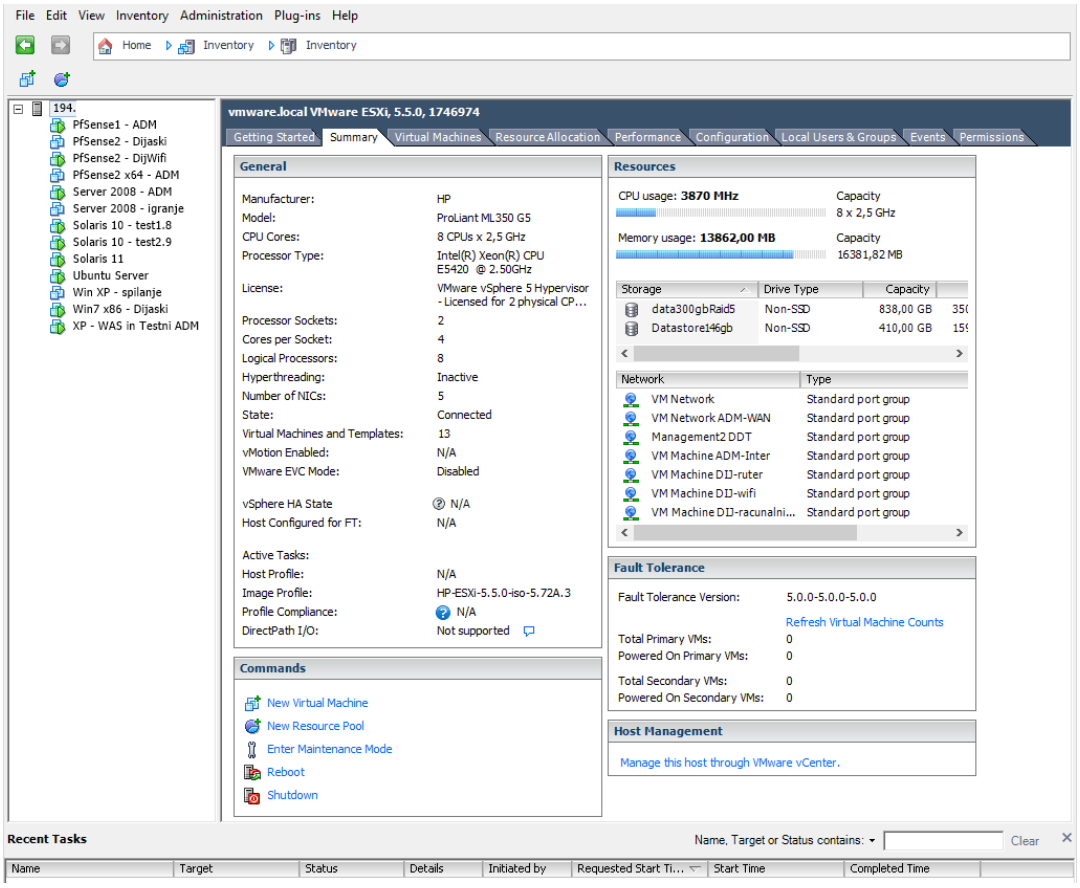
Najbolj znani programi, ki delujejo neposredno na strojni opremi so VMware vSphere Hypervisor (ESXi), Microsoft Hyper-V in Citrix XenServer. Pri licenciranju programske opreme v virtualnem okolju veljajo ista pravila, kot če bi programsko opremo namestili na fizični računalnik.

Slika 11: Bare metal hypervisor (vir: vmware.com)

Na področju virtualizacije je podjetje VMware vodilni svetovni ponudnik programske opreme. Ponuja tako brezplačne licence kot tudi zelo drage za hipervizorje namenjene podatkovnim skladiščem. Pri nas smo se odločili za brezplačno različico, ki teče neposredno na strojni opremi (*bare metal*), VMware vSphere ESXi v arhitekturi x64 [12]. ESXi, kot brezplačna verzija ESX, je ta trg prišel v letu 2007, konec leta 2012 pa se je VMware odločil za preimenovanje produkta v vSphere Hypervisor in s tem postopno ukinitvijo plačljive ESX različice. Trenutno je zadnja aktualna verzija 5.5. V zadnjih brezplačnih verzijah so pri zahtevah za strojno opremo umaknili omejitev na 32 GB spomina, omogočili pa so neomejeno število fizičnih procesorjev in njegovih jeder na gostitelju (*angl. host*). Ohranili pa so omejitev osem virtualnih procesorjev na posamezno virtualno napravo. Upravljanje hipervizorja, konfiguriranje virtualk in dostop do njih poteka preko drugega računalnika v omrežju oz. internetu, s pomočjo programa vSphere Client (*slika 12*), s prihajajočimi verzijami pa se bo upravljanje preselilo na vSphere Web Client.

Na *slika 12* je jasno razvidno, kako velik prihranek nam omogoča virtualizacija. Na samo enem računalniku imamo namreč nameščenih 13 operacijskih sistemov, medtem ko bi brez virtualizacije potrebovali kar 13 fizičnih računalnikov, poraba električne energije pa bi bila

neprimerno višja. Produkciji je namenjenih pet virtualnih računalnikov, druge pa so namenjene testiranju ter analizi prometa za potrebe te naloge.

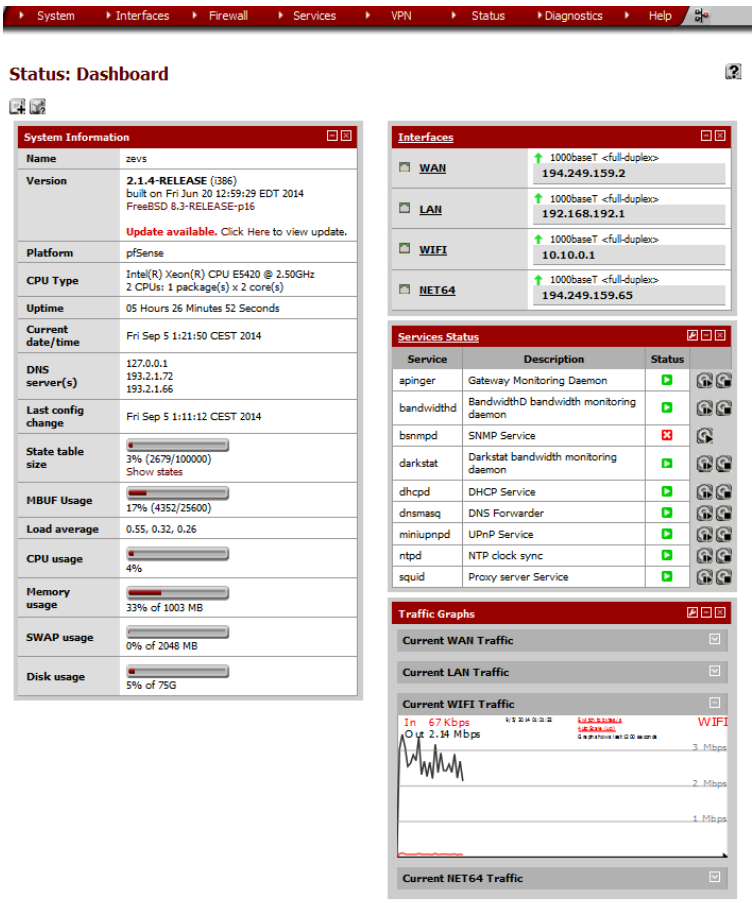


Slika 12: vSphere Client - program za upravljanje z virtualkami

2.4.2 pfSense

Ključni program za dijaški dom, ki teče na hipervizorju, je odprtokodni programski usmerjevalnik in požarni zid pfSense. Razvit je na operacijskem sistemu FreeBSD, ki so ga prikrojili za uporabo kot požarni zid in usmerjevalnik, vse skupaj pa zapakirali v ličen spletni vmesnik preko katerega ga tudi upravljamo (*slika 13*). Projekt pfSense se je iz takratnega projekta M0n0wall – požarni zid začel ločeno razvijati leta 2004. M0n0wall je bil osredotočen bolj na namensko vgrajeno strojno opremo (*embedded hardware*) s tem ko so pfSense razvijali za namestitve na poljuben računalnik. Tako PfSense že ob namestitvi ponuja ogromno funkcij, ki pa jih lahko zaradi odprtokodnosti že z enim miškinim klikom še dodatno nadgradimo z raznimi vnaprej pripravljenimi paketi drugih proizvajalcev (popularen je recimo paket Squid). PfSense je po svetu zelo razširjen in izjemno priljubljen, zato ogromno podpore

dobimo tudi na raznih forumih in *mailing* listah, prav tako pa nam je na voljo tudi komercialna podpora [13, 14].



Slika 13: pfSense spletni vmesnik za upravljanje

Pri nas imamo nameščena dva pfSense virtualna sistema, torej za vsako omrežje svojega. To se nam je zdelo smiselno, saj pfSense uporabljamo kot usmerjevalnik (NAT translacija), v primeru vzdrževalnih del ali odpravljanja napak pa je posledično prizadeto samo dijaško ali samo administrativno omrežje. Funkcionalnosti, ki jih uporabljamo v našem sistemu, so:

- DHCP (*Dynamic Host Configuration Protocol*) strežnik – avtomatsko dodeljevanje IP naslovov napravam, na podlagi vnaprej vnešenih MAC naslovov.
- Požarni zid (*angl. firewall*) – služi za omejevanje in zaščito uporabnikov med našim in drugimi omrežji. Uporabnikom omogočamo dostop do vseh standardnih vrat (*angl. port*) zunanjih IP-jev, nestandardne (nad 1024) pa blokiramo. V primeru posameznih potreb omogočimo tudi druge dostope. Do letošnjega šolskega leta smo, zaradi

premalo dostopovnih točk in zgolj 54 Mb/s prenosa na posamezno točko, uporabnikom onemogočali dostop do P2P prometa (torenti), v letošnjem letu ob nakupu novih, zmogljivejših točk bomo to omejitev poskusno odpravili.

- preslikava omrežnih naslovov (*NAT*) za omogočanje skupne rabe interentne povezave,
- Navidezno privatno omrežje (*VPN*) – s pomočjo OpenVPN orodja, ki ga že v osnovi ponuja pfSense imamo urejen oddaljen dostop do internega omrežja.
- Urejevalnik prometa (*angl. traffic shaper*) – s pomočjo tega orodja lahko omejujemo oz. oblikujemo promet, ki gre skozi pfSense. Z njim dijakom določimo download in upload hitrost na 5 Mb/s (zaradi ozkega grla pri dostopovnih točkah), priotiriziramo promet za nekatere protokole (MSRDP, VNC, MSN, IMAP, POP3, DNS) in nekatere spletne igre.
- Poročanje in nadzor nad sredstvi (*angl. resources*) – s pomočjo RRD grafov lahko spremljamo trenutno stanje in zgodovino številnih sredstev. To orodje bom uporabljal tudi pri nadaljnji analizi omrežja.
- Proxy strežnik – dodatno nameščen paket Squid, ki deluje kot posrednik med zahtevami uporabnikov do drugih strežnikov.
- Beleženje dnevnikov na oddaljen Syslog strežnik.

2.4.3 Microsoft Windows Server 2008 Standard x64

Ravno ob nakupu novega strežnika je na trg prišla nova različica strežniškega operacijskega sistema Windows Server 2008. Odločitev za prehod iz poznane starejše verzije Windows Server 2003 na novo ni bila težka, saj ni smiselno nameščati starejših edicij, če imamo dostop do novejših. Pri namestitvi standard edicije, smo morali upoštevati omejitev podpore do štirih procesorjev in možnost uporabe do 32 GB spomina pri x64 verziji. Uporabo operacijskega sistema z letnim najemom nam je omogočilo Ministrstvo za šolstvo in šport. Ta platforma v našem dijaškem domu predstavlja osnovo za delo večine zaposlenih, saj omogoča izvajanje številnih poslovnih aplikacij, vse od računovodske do aplikacije za vodenje evidence dijakov in študentov. Tako MS Server 2008 našemu poslovnemu okolju omogoča sledeče funkcionalnosti:

- Aktivni imenik oz. domenski strežnik (*Active directory*, krajše AD) – omogoča upravljanje z uporabniki (zaposlenimi), datotekami, tiskalniki, računalniki itd. Z njim pridobimo enotno prijavo uporabnikov na vseh računalnikih v isti domeni in s tem dostop do istih podatkov.

- DNS (*angl. Domain Name Server*) strežnik – potreben samo za delovanje domenskega strežnika, saj za primarni DNS strežnik uporabljamo pfSense.
- Podatkovni strežnik – omogoča deljenje datotek v omrežju.
- Tiskalniški strežnik – omogoča uporabo vseh tiskalnikov v domu na podlagi določenih pravil.
- Aplikacijski strežnik – z njim distribuiramo razne programe oz. želene popravke do uporabnikov.
- Strežnik za popravke (*Windows Server Update Services*, krajše WSUS) – popravke in servisne pakete, ki jih izda Microsoft shrani v bazo in jih ob določenih urah oz. dneh razpošlje drugim računalnikom.
- SQL strežnik – strežnik za podatkovne baze.

Poleg teh funkcionalnosti, pa so na njem nameščeni namenski programi oz. podatkovne baze, brez katerih bi bilo delo zaposlenih zelo težavno:

- SAOP – program za vodenje računovodstva,
- SEZAM – informacijski sistem strokovnih delavcev za celovito vodenje evidence dijakov (podatkov o dijakih, evidenca koriščenja obrokov, pisanje e-dnevnika, evidentiranje nočnih prihodov v dom, obračunavanje mesečne oskrbnine itd.),
- Hostel – program za vodenje turistične dejavnosti v poletnem času,
- Špica – evidentiranje prisotnosti zaposlenih.

2.4.4 Drugi operacijski sistemi

V VMware virtualnem okolju imamo nameščenih tudi nekaj drugih operacijskih sistemov. Nameščeni so z namenom testiranja sprememb in samega delovanja celotnega sistema, pridobivanja podatkov za to diplomsko delo, nadzor Ubiquiti dostopovnih točk, ipd. Za potrebe tega diplomskega dela smo namestili operacijske sisteme Solaris 10, Solaris 11 in Ubuntu. Microsoft Windows 7 je namenjen nadzoru dostopovnih točk in uporabi v dijaškem omrežju, MS Windows XP pa služi za nadzor administrativnega omrežja ter kot notranji spletni strežnik za prikaz prisotnosti zaposlenih (ta aplikacija ni na Windows Server 2008 strežniku, ker ga naša verzija Špica programa za evidentiranje zaposlenih ne podpira v celoti).

3 METODOLOGIJA

3.1 Cilji meritev

V dijaškem domu se vrstita dve skupini uporabnikov. V poletnem času, ko so dijaki na zasluženih počitnicah, gostimo popotnike iz celega sveta, saj delujemo kot hostel. Naši dijaki so se zadnje leto občasno pritoževali, da jim brezžično omrežje deluje počasi in nestabilno. Vedeli smo, da za to ni kriva zgolj 100 megabitna linija, vendar smo kljub temu storili prvi korak k reševanju problema v tej smeri in tako nadgradili linijo na 1 Gb/s. Večji problem namreč predstavljajo razporeditev ter zastarelost dostopovnih točk iz leta 2009 ter vse več aktivnih naprav v omrežju z različno kvalitetnimi mrežnimi karticami.

Kot pomoč pri reševanju teh težav smo se odločili, da bomo v prvem sklopu analiz navade uporabnikov – najprej tujcev v mesecu avgustu, nato pa še navade dijakov v prvem tednu šolskega leta. Ob tem bomo poskušali ugotoviti, kdaj so naši uporabniki najbolj aktivni, katere dostopovne točke so najbolj obremenjene ter do katerih strani oz. IP-jev ter preko katerih protokolov oz. vrat (*angl. port*) največkrat dostopajo do spletnih vsebin. Na ta vprašanja bomo poskušali odgovorit s pomočjo več programov oziroma paketov znotraj pfSense okolja ter s pomočjo programskega vmesnika za dostopovne točke. V drugem sklopu analiz se bomo osredotočili na mednarodne povezave, pri katerih nas bo zanimala hitrost prenosov ob različnih parametrih. V ta namen bomo na različnih operacijskih sistemih (tudi znotraj virtualnega okolja VMware), z različnimi parametri TCP nastavitev, s pomočjo skript poganjali program *wget* ter rezultate kasneje obdelali v analitičnem spletnem orodju MicroStrategy Analytics Express.

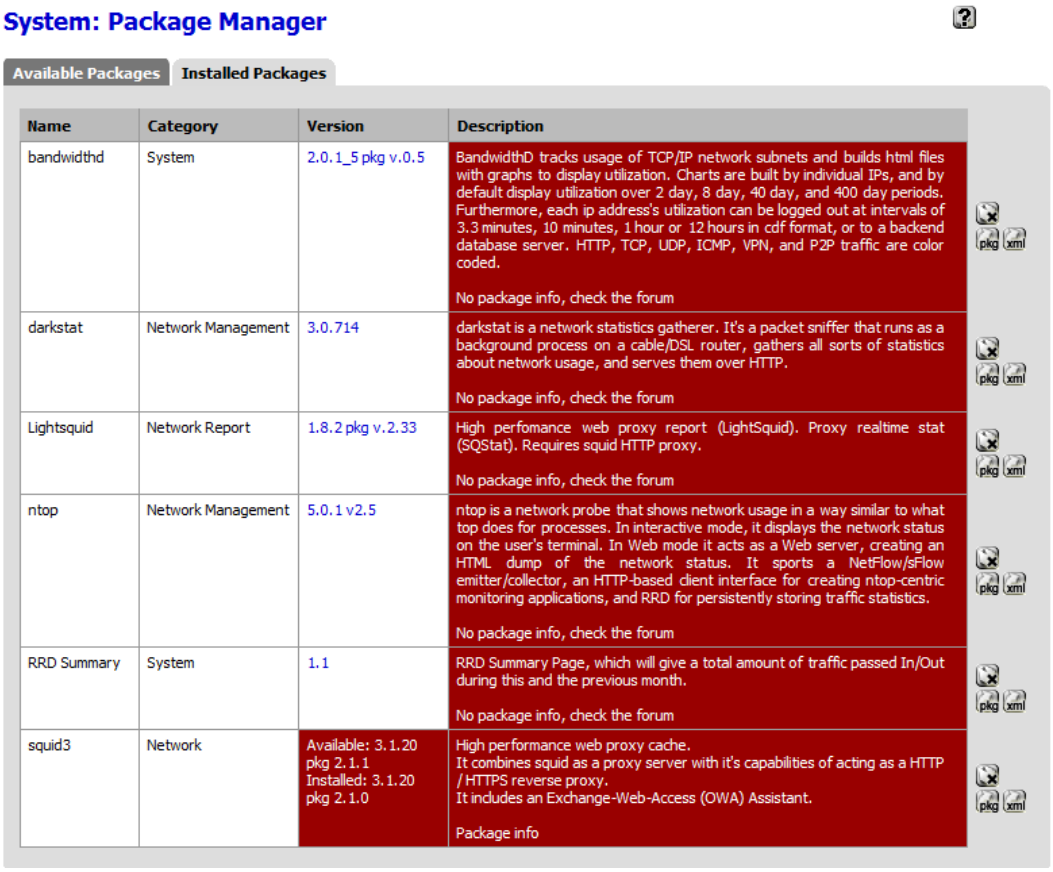
S pomočjo teh statistik bomo dobili jasnejši vpogled v uporabo našega omrežja in ga tako v tem šolskem letu poskušali optimizirati.

3.2 Opis meritev uporabe omrežja

Za prvi sklop meritev smo si v že postavljenem virtualnem okolju za nalogo najprej zadali postavitev novega pfSense sistema. V hipervizorju smo za to virtualno napravo izbrali vse potrebne nastavitve. Odločili smo se za x86 FreeBSD platformo, saj ne bomo potrebovali več kot 4 GB spomina, nastavili smo en virtualen procesor, 2 GB spomina ter 60 GB prostora v podatkovni shrambi. Izbrali smo tudi potrebno število mrežnih kartic ter se lotili same namestitve. Po koncu nameščanja programa nastavimo potrebne IP naslove na posameznih karticah, od tu dalje pa poteka upravljanje preko spletnega vmesnika. Ko smo nastavili DNS strežnike in ustrezen prehod (*angl. gateway*) na WAN vmesniku, smo dobili možnost nameščanja dodatnih paketov. Za potrebe nadzora in kasnejših analiz smo se odločili za uporabo sledečih paketov (*slika 14*):

- BandwidthD – generira lahko HTML stran, ki nam omogoča vpogled v uporabo različnih protokolov v omrežju ter grafičen prikaz statistik, kot so dnevna, tedenska, mesečna ter letna uporaba določenih protokolov, hitrosti prenosov in podobno. Paket z enim klikom namestimo, potem je naša naloga le, da mu v nastavitvah povemo na katerem mrežnem vmesniku naj posluša promet, kam naj shranjuje podatke ter omogočimo izrisovanje grafov.
- Squid 3 – je posredniški (*angl. proxy*) strežnik, ki deluje kot posrednik med brskalnikom uporabnika in internetom. Uporabnik od njega zahteva neko stran, strežnik jo poišče, shrani v svojo bazo ter jo posreduje uporabniku.
- LightSquid – kot spletno stran prikaže osnovno statistiko squid proxy strežnika. Kot vir podatkov uporablja lokalno podatkovno bazo od nameščenega squid proxy strežnika.
- RRD (*angl. Round - Robin Database*) graphs – zelo zmogljivo, že vnaprej vgrajeno, orodje za spremljanje raznoraznih statistik pfSense sistema, med drugim nam bo služilo tudi za prikaz uporabe omrežja v določenem obdobju.
- Firewall – požarni zid je osnovna funkcionalnost pfSense sistema. S pomočjo primerno nastavljenih pravil ter beleženja le-teh bomo nadzirali in analizirali uporabo omrežja.

S pomočjo sprotnega beleženja števila uporabnikov starih dostopovnih točk ter grafičnega prikaza uporabe omrežja v nadzornem sistemu novih dostopovnih točk, bomo prikazali kdaj uporabniki največ dostopajo do spletnih vsebin.



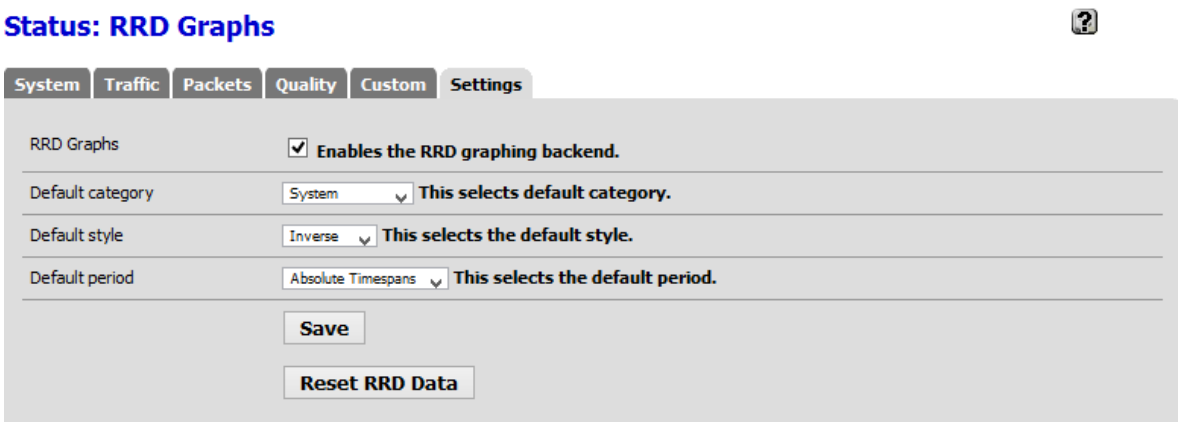
Slika 14: Nameščeni dodatni paketi v pfSense

Potrebne podatke za analize smo po koncu beleženja iz standardnih formatov beleženja pretvorili v nam bolj pregledne ter jih s pomočjo skripte shranili v *CSV* datoteke. Kasneje se je izkazalo, da teh datotek zaradi preobilice podatkov oz. prevelikih datotek, ne bo moč analizirati v programu Microsoft Excel, saj se ob odpiranju tako velike datoteke program sesuje. Odločili smo se, da bomo na računalnik namestili podatkovno zbirko PostgreSQL, in s pomočjo pripadajočega orodja pgAdmin uvozili že prej omenjene *CSV* datoteke. Izbor podatkov potrebnih za posamezno analizo bomo pridobili s pomočjo SQL (*angl. Structured Query Language*) poizvedb, rezultate poizvedb pa bomo obdelali s programom Microsoft Excel.

3.2.1 RRD grafi

PfSense v ozadju neprekinjeno beleži informacije o stanju sistema, ter kasneje s pomočjo odprtokodne aplikacije RRD toolset, le-te grafično predstavi. V formatu RRD graph nam

lahko analizira in prikaže sledeče informacije (*slika 15*): *System* (prikaz obremenjenost procesorja, stanje spomina, itd.), *Traffic* (analiza prometa na vseh omrežnih vmesnikih, odhodni/dohodni promet, OpenVPN promet, itd.), *Packets* (analiza paketov na vseh omrežnih vmesnikih) in *Quality* (prikaže morebitno izgubo paketov oz. kvaliteto povezave na prehodih - *gateways*). Privzeto nam prikazuje grafe za vnaprej določena obdobja, s pomočjo zavihka *Custom* pa lahko tudi sami izberemo obdobje. Pri analizi se bomo tudi mi posluževali te opcije. [15]



Slika 15: Vmesnik za nastavitve in prikaz RRD grafov

3.2.2 Proxy strežnik Squid

V urejevalniku paketov smo program najprej namestili. Po nameščanju se nam v meniju pojavi možnost upravljanja tega paketa oz. programa. V nastavitvah smo omogočili transparentno (*angl. transperent proxy*) delovanje, tako da uporabniku v brskalniku ni potrebno ročno nastavljati podatkov o strežniku. Določili smo mu na katerem vmesniku naj posluša promet (*WIFI* vmesnik), kam ter za koliko dni naj na lokalni disk shranjuje podatke (odločili smo se za 60 dni, saj je ta strežnik zelo potraten pri uporabi prostora). Slabost zbiranja podatkov na ta način je, da nimamo vpogleda v HTTPS promet, saj le ta potuje po zaščiteni poti in se kot tak ne beleži. Težavo smo odpravili s pomočjo beleženja pravil v požarnem zidu samega pfSense sistema. Pri beleženju nezaščitenega prometa imamo prednost v tem, da vidimo celoten ciljni naslov in ne zgolj IP-ja kot je to pri beleženju v požarnem zidu. Primer vrstice v squid *access.log* datoteki:

1407362402.933 721 10.10.4.87 TCP_MISS/200 4143 GET http://www.booking.com/srcompset.es.html? - DIRECT/5.57.16.220 text/html

S pomočjo *python* skripte (*koda 1*) in uporabe regularnih izrazov (*RegEX*), smo v *CSV* datoteko zapisali le ključne podatke za našo analizo (*slika 16*).

ts	domain	unixtime	duration	source_ip	bytes	request	url
2014-08-28 00:00:02	cedexis.com	1409176802.198	1110	10.10.4.73	443	GET	http://sin1.voxcloud.cedexis.com/r16/r16.js?
2014-08-28 00:00:02	domain-kb.com	1409176802.208	1111	10.10.3.109	17126	GET	http://domain-kb.com/static/style/bootstrap/css/bootstrap.min.css

destination_ip	mime
107.6.125.18	application/x-javascript
192.95.22.46	text/css

Slika 16: Podatki iz obdelave access.log datoteke

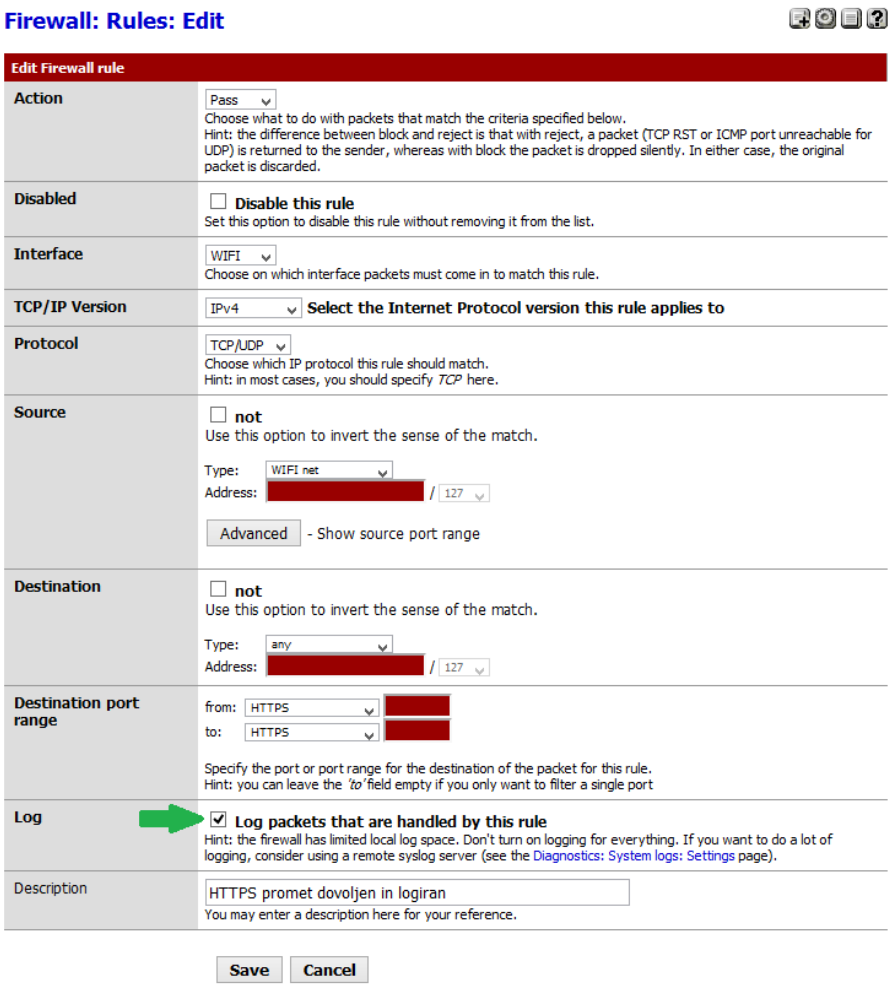
```
import argparse
import re

parser = argparse.ArgumentParser(description='Extract data from Squid log to CSV file.')
parser.add_argument('source_filename', type=str, help='Input Squid LOG file')
parser.add_argument('destination_filename', type=str, help='Output CSV file')

args = parser.parse_args()

with open(args.source_filename, "r") as source_file:
    with open(args.destination_filename, "w") as destination_file:
        for line in source_file:
            replaced_line = re.sub(
                r"(\S+)\s+(\d+)\s+(\S+)\s(?!NONE/400)\S+\s(\d+)\s(\S+)\s(\S+)\s-\s\S+/(\S+)\s(?=\S+/\S)(\S+)",
                r"\1\t\2\t\3\t\4\t\5\t\6\t\7\t\8",
                line,
                0,
                re.IGNORECASE
            )
            destination_file.write(replaced_line)
```

Koda 1: Skripta za izvoz podatkov v CSV datoteko

3.2.3 Požarni zid pfSense

Zgrajen je na osnovi zelo močnega pf *(angl. Packet Filter*) požarnega zida iz OpenBSD operacijskega sistema [16]. Požarni zid deluje tako, da gre celoten promet skozi seznam definiranih pravil, če paket ustreza določenemu pravilu ga požarni zid bodisi spusti skozi ali pa zavrne. Vsa pravila se izvajajo od vrha proti dnu seznama, ko paket naleti na ustrezno pravilo se le to izvrši, ostala pravila pa preskoči. Pravila se pišejo s pomočjo preglednega vmesnika, ki nam omogoča ogromno nastavitev oziroma določitev pogojev izvrševanja le-teh.

Naša analiza bo temeljila na beleženju definiranih pravil, to omogočimo tako, da postavimo kljukico pri funkciji *Log packets* (*slika 17*). Na ta način beležimo uporabo sledečih protokolov: HTTP, HTTPS, FTP, SSH, IMAP, IMAP/S, SMTP/S, POP3 in POP3/S.



Slika 17: Primer kreiranja pravila v požarnem zidu

PfSense smo nastavili tako, da vsa beleženja aktiviranih pravil pošilja v oddaljen *Syslog* strežnik. Za ta namen smo ga vzpostavili v NAS (*angl. Network-attached storage*) strežniku Qnap TS-879 Pro. *Syslog* predstavlja standard v računalništvu za beleženje sporočil. Sporočila, ki se beležijo iz različnih komponent sistema imajo svoje oznake, tako da jih lažje razločimo. Na isti *syslog* strežnik lahko beležimo sporočila iz več mrežnih naprav, če tak način beleženja podpirajo [17].

Primer zapisa v log datoteki na *syslog* strežniku:

<134>1 2014-08-07T22:38:50+02:00 194.249.X.Y pf - - - pf: 10.10.4.129.52061 > 213.150.2.216.443: Flags [S], cksum 0x5b91 (correct), seq 506296969, win 65535, options [mss 1460,nop,wscale 4,nop,nop,TS val 1045581906 ecr 0,sackOK,eol], length 0

<134>1 2014-08-07T22:38:50+02:00 194.249.X.Y pf - - - pf: 00:00:00.051285 rule 133/0(match): pass in on em2: (tos 0x0, ttl 64, id 16605, offset 0, flags [DF], proto TCP (6), length 64)

Izmed množice podatkov v posameznem zapisu se bomo osredotočili le na bistvene podatke za našo analizo: datum in uro dostopa, IP od prehoda, notranji izvorni IP in port, zunanji ciljni IP in port (*slika 18*). Podatke v *CSV* datoteko izvozimo z uporabo skripte predstavljene v kodi (*koda 2*).

date	hour	gateway	protocol	source_ip	source_port	destination_ip	destination_port
2014-08-08	04:33:43	194.249.159.2	TCP	10.10.3.216	49377	173.252.74.27	443
2014-08-08	04:33:43	194.249.159.2	TCP	10.10.3.216	49378	31.13.93.33	443
2014-08-08	04:33:43	194.249.159.2	TCP	10.10.3.216	49379	69.171.245.49	443

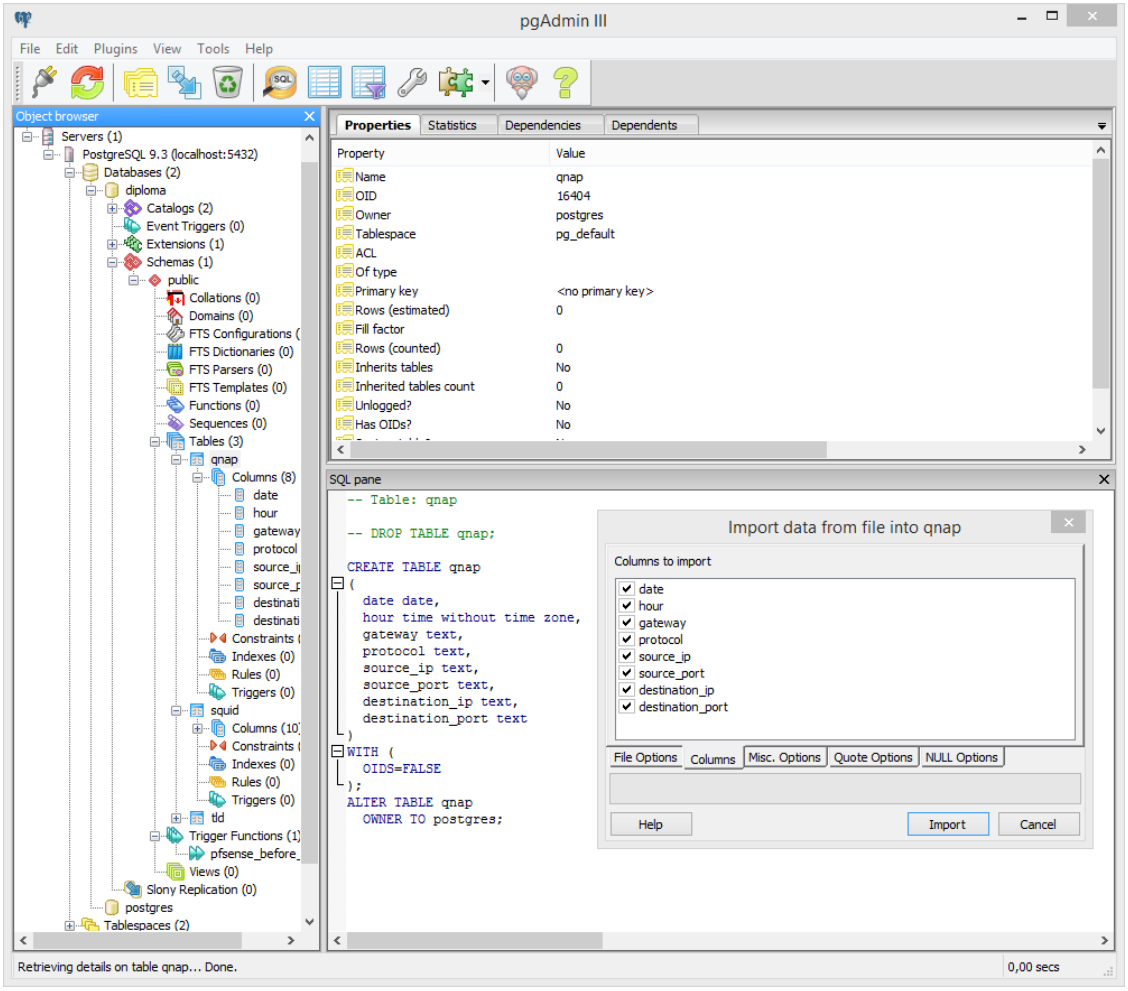
Slika 18: Izbor podatkov beleženja požarnega zidu

```
import argparse
import re

parser = argparse.ArgumentParser(description='Extract data from Qnap firewall log to CSV file.')
parser.add_argument('source_filename', type=str, help='Input Qnap firewall LOG file')
parser.add_argument('destination_filename', type=str, help='Output CSV file')

args = parser.parse_args()

with open(args.source_filename, "r") as source_file:
    with open(args.destination_filename, "w") as destination_file:
        first_line = source_file.readline()
        while True:
            if not first_line: break
            second_line = source_file.readline()
            if not second_line: break
            line = first_line + second_line
            replaced_line = re.sub(
                r"<134>1\s(\S+)T(\S+)\+\S+\s(\S+).+proto\s(\S+).+\n" +
                    r".+pf:\s+((?:\d{1,3}\.){3}\d{1,3})\.(\d+)\s>\s((?:\d{1,3}\.){3}\d{1,3})\.(\d+):.+",
                r"\1\t\2\t\3\t\4\t\5\t\6\t\7\t\8",
                line,
                0,
                re.IGNORECASE
            )
            if replaced_line != line:
                destination_file.write(replaced_line)
            first_line = second_line
```

Koda 2: Skripta za izvoz podatkov v CSV datoteko

3.2.4 Uvoz v podatkovno bazo PostgreSQL

Ko smo imeli datoteki s podatki pripravljeni, smo z orodjem *pgAdmin* kreirali podatkovno bazo. V njej smo ustvarili dve tabeli, eno za qnap-firewall ter eno za squid-pfsense podatke. Posamezno datoteko oziroma njene podatke smo uvozili s pomočjo funkcije *Import*, v kateri smo izbrali katera polja oz. stolpce iz posamezne *CSV* datoteke bi radi uvozili (*slika 19*). Zaporedje podatkov v *CSV* datotekah, ločenih s tabulatorjem, mora bit v enakem zaporedju kot so stolpci v posamezni tabeli.



Slika 19: Prikaz vmesnika pgAdmin in uvoz qnap-firewall csv datoteke

3.3 Opis meritev mednarodnih prenosov z različnimi TCP nastavitvami

V drugem delu meritev bomo merili hitrosti prenosov iz strežnikov v državah na različnih kontinentih (*tabela 1*). Meritve smo opravljali tri tedne, vendar se bomo zaradi preobilice podatkov in časovne stiske osredotočili le na dva tedna (25.8. – 7.9.2014). Da bodo meritve merodajne, smo na internetu poiskali strežnike, ki imajo možnost prenosa enake datoteke, naleteli smo na datoteko *Fedora-i386-20-20131211.1-sda.raw.xz* v velikosti 114 MB. Hitrost prenosa je odvisna predvsem od mrežnih povezav in kapacitet do posamezne lokacije, a kot bomo videli kasneje, je odvisna tudi od nastavitev oz. implementacije TCP/IP mrežnega protokola v samem operacijskem sistemu. Da smo lahko analizirali te razlike, smo postavili šest operacijskih sistemov (*tabela 2*).

Država	Koda države	Hostname	Hitrost priključka na internet	Latenca
Argentina	AR	fedora.xfree.com.ar	100 Mb/s	260.5 ms
Avstralija	AU	mirror.optus.net	2 Gb/s	324.9 ms
Brazilija	BR	www.las.ic.unicamp.br	1 Gb/s	235.7 ms
Hong Kong	HK	ftp.cuhk.edu.hk	1 Gb/s	312.8 ms
Kanada	CA	mirror.csclub.uwaterloo.ca	200 Mb/s	135.8 ms
Singapur	SG	mirror.nus.edu.sg	1 Gb/s	240.2 ms
Velika Britanija	GB	www.mirrorservice.org	8 Gb/s	46.3 ms
ZDA	US	mirrors.tummy.com	90 Mb/s	151.2 ms

Tabela 1: Prikaz držav, strežnikov in hitrosti povezav

Operacijski sistem	Oznaka	Lokacija OS	TCP nastavitve
Mac OS X 10.9	MacPro-OSx	Računalnik MacPro	Privzete
Solaris 10	Solaris10-ip8	VMware ESXi	Optimirane
Solaris 10	Solaris10-ip9	VMware ESXi	Privzete
Solaris 11	Solaris11-ip13	VMware ESXi	Privzete
Solaris 11	Solaris11-wifi	Računalnik Intel i5	Privzete
Ubuntu	Ubuntu	VMware ESXi	Privzete

Tabela 2: Seznam operacijskih sistemov

Da bi čim bolj izločili zunanje dejavnike, ki bi vplivali na hitrost prenosov, so bili vsi računalniki priključeni na isto stikalo in 1 Gb/s linijo. Izjema je bila le brezžična povezava na računalniku, ki je bil priključen na UAP-Pro dostopovno točko. Za prikaz pomembnosti TCP nastavitev smo namestili dva identična operacijska sistema Solaris 10 z različnimi mrežnimi nastavitvami. Privzete nastavitve v Solaris 10 so za današnje čase neustrezne, kar se pozna predvsem na povezavah z visoko latenco, na tako imenovanih *high bandwith-delay product paths* [18]. TCP/IP nastavitve smo optimizirali na Solaris10-ip8 sistemu z naslednjimi ukazi [19, 20, 21, 22, 23]:

```
ndd -set /dev/tcp tcp_max_buf 16777216
ndd -set /dev/tcp tcp_cwnd_max 8388608
ndd -set /dev/tcp tcp_recv_hiwat 4194304
ndd -set /dev/tcp tcp_xmit_hiwat 4194304
```

Privzete vrednosti zgornjih parametrov smo na sistemu Solaris10-ip9 pustili nedotaknjene, in sicer: *tcp_max_buf* - 1048576, *tcp_cwnd_max* - 1048576, *tcp_recv_hiwat* - 49152, *tcp_xmit_hiwat* - 49152. V novejšem sistemu Solaris11 so te neoptimalne nastavitve popravili.

3.3.1 Prenos datotek s programom *wget*

Na vseh računalnikih smo namestili program *wget*, ki nam bo služil za prenašanje datotek s pomočjo pripravljene skripte. *Wget* je brezplačno odprtokodno orodje, ki nam preko protokolov HTTP, HTTPS in FTP omogoča prenos datotek iz spleta kar v ukazni vrstici. Za njegovo uporabo smo se odločili, ker je podprt na večini operacijskih sistemov ter na vseh deluje enako (z enakimi parametri) [24]. Beleženje prenosa datoteke smo preusmerili v dnevnik s sledečim ukazom:

```
wget --progress=dot -r $1 2>&1 > "logs/wget-$2-$(date +%y%m%d_%H%M%S).out"
```

Za potrebe analize pa nam samo privzeto beleženje povprečne hitrosti ni zadostovalo, saj smo potrebovali tudi točen čas prenosa posameznega bloka. To smo naredili z dodatno skripto, v kateri smo pred vsako zabeleženo vrstico dodali še časovni žig (angl. *timestamp*).

Skripta *wget-log.sh*, ki nam omogoča beleženje imena strežnika s časovnim žigom, lokacijo datoteke podamo kot parameter skripte, vsaka zabeležena vrstica v dnevniku (le-ta predstavlja velikost bloka 50 KB) ima svoj časovni žig (*koda 3*):

```
wget --progress=dot -r $1 2>&1 |
while read out; do
  echo "$(date '+%d/%m/%y %H:%M:%S'): $out";
done > "logs/wget-$2-$(date +%y%m%d_%H%M%S).out"
```

Koda 3: Skripta za beleženje časovnega žiga prenosa posameznega bloka

Primer zapisa v dnevnik s časovnim žigom:

```
02/09/14 15:00:00: --2014-09-02 15:00:00--  http://mirror.csclub.uwaterloo.ca/fedora/linux/20/Images/i386/...
02/09/14 15:00:00: Resolving mirror.csclub.uwaterloo.ca. 129.97.134.71
02/09/14 15:00:00: Connecting to 129.97.134.71|:80... connected.
02/09/14 15:00:00: HTTP request sent, awaiting response... 200 OK
02/09/14 15:00:00: Length: 119314240 (114M) [text/plain]
02/09/14 15:00:00: Saving to: `mirror.csclub.uwaterloo.ca/…/Images/i386/Fedora-i386-20-20131211.1-sda.raw.xz'
02/09/14 15:00:00:
02/09/14 15:00:01: 0K ..... ..... .....  0%  205K 9m28s
02/09/14 15:00:01: 50K ..... ..... .....  0%  400K 7m9s
02/09/14 15:00:01: 100K ..... ..... .....  0% 11.5M 4m49s
...
02/09/14 15:06:03: 116500K ..... .....              100% 2.22M=6m2s
02/09/14 15:06:03: 2014-09-02 15:06:03 (322 KB/s) - `mirror.ca/…/Fedora-i386-20-20131211.1-sda.raw.xz' saved
[119314240/119314240]
02/09/14 15:06:03:
02/09/14 15:06:03: FINISHED --2014-09-02 15:06:03--
02/09/14 15:06:03: Downloaded: 1 files, 114M in 6m 2s (322 KB/s)
```

Skripta *wget-batch.sh,* pokliče skripto *wget-log.sh*, ki ji preko parametrov poda URL datoteke za prenos in ime strežnika, se samodejno poganja vsake dve uri s pomočjo *cron job* ukaza:

```
matjaz@solaris:~$ crontab -e
0 1,3,5,7,9,11,13,15,17,19,21,23 * * * /home/matjaz/wget-batch.sh
```

Primer ene vrstice v skripti *wget-batch.sh*:

```
./wget-log.sh http://fedora.xfree.com.ar/linux/releases/20/Images/i386/Fedora-i386-20-20131211.1-sda.raw.xz
Argentina_100mb
```

3.3.2 Razčlenjevanje dnevnikov

Dnevnik, ki ga dobimo s poganjanjem *wget* ukaza, ni primeren za nadaljnje obdelave v analitičnih programih. V ta namen smo si pripravili *python* skripto, ki nam generira lepo strukturirane *CSV* datoteke. Za obdelavo teh datotek smo našli kakovostno spletno analitično orodje MicroStrategy Analytics [25], ki nam že omogoča uvoz *CSV* datotek. Se je pa kasneje izkazalo za zelo omejujoče, saj nam dovolijo le datoteke v velikosti 200 MB, ter uporabniku namenijo skupno 1 GB prostora. Kljub prošnjam preko elektronske pošte, da nam začasno odstranijo te omejitve, nam niso ugodili. Zato smo bili primorani spreminjati *python* skripto, da smo izločili nepotrebne podatke oziroma nekatere preoblikovali v manj znakov. Za nepotrebne se je pri analizi izkazal predvsem URL.

Datoteka *CSV,* ki jo dobimo iz razčlenjevanja dnevnika po vrsticah (*koda 4* in *koda 5*), vsebuje za vsak blok (blok je velikosti 50 KB) sledeče zapise:

- *computer* – ime računalnika na katerem se je *wget* izvajal
- *country* – država kjer se nahaja strežnik
- *ft_start_month* – mesec izvedbe prenosa
- *ft_start_day* – dan izvedbe prenosa
- *ft_start_dayofweek* – dan v tednu začetka prenosa
- *ft_start_hour* – ura v dnevu začetka prenosa
- *block_nr* – zaporedna številka 50 KB velikega bloka
- *block_avg_speed* – povprečna hitrost prenosa posameznega bloka
- *block_elapsed_ti*me – potreben čas za prenos posameznega bloka

Primer zapisa v končni verziji *CSV* datoteke, za računalnik *Solaris10-ip8*:

```
block_elapsed_time,ft_start_day,country,block_avg_speed,ft_start_month,file_avg_speed,file_elapsed_time,ft_start_hour,co
mputer,block_nr,ft_start_dayofweek
0.33,25,AR,150.0,08,45.6,30.1,1,solaris10-ip8,1,Mon
0.33,25,AR,150.0,08,45.6,30.1,1,solaris10-ip8,2,Mon …
```

Naslednja težava pri MicroStrategy orodju je bila, da ni bilo mogoče uvoziti podatkovnega tipa *timestamp.* Poizkušali smo več formatov YYYY/MM/DD HH:MI:SS, DD/MM/YYYY, MM/DD/YYYY, itd. vendar neuspešno. To smo rešili tako, da smo v *CSV* datoteko ločeno zapisali vsak atribut časovnega žiga (*koda 4*).

Ko smo podatke že uspešno prikazali na grafu, smo naleteli na naslednjo oviro, specifično bolj za naše podatke. Orodje namreč nima nobene nastavitve, da bi ustrezno prilagodili X os, in s tem naš prepodroben graf prikazali na enem zaslonu. V ta namen smo *python* skripto prilagodili, da je zapisovala povprečja za vsakih 15 blokov, kar se je izkazalo ravno dovolj za prikaz celotnega grafa na enem zaslonu. (*koda 5*)

Ob poganjanju *python* skripte, se je izkazalo, da je v nekaterih dnevnikih kar precej napak, ki so onemogočili izvajanje skripte. Te težave smo sproti reševali, saj skripta ni predvidevala odstopanja od vnaprej določene strukture dnevnika.

```
with open(logfile, "r") as infile:
    for line in infile:
        # Zacetek parsanja
        i = i + 1
        split_line = line.split()
        if i == 1:
            # Prva vrstica je URL
            split_line = line.split()
            url = split_line[4]
        if i == 2:
            # Druga vrstica je IP
            ip = split_line[len(split_line)-1]
            datablock['country'] = country[ip]
        if "Length:" in split_line:
            # Beremo dokler ne zaznamo stringa Length in si zapisemo offset
            length = int(split_line[3])
            offset = i - 5
            continue
        # Vrstice med vrstico z Length in vrstico 8+offset lengtha ignoriramo.
        if i < offset + 8: continue
                                        if "........." in split_line:
                                                # Podatkovna vrstica
                                                start_tr = True

                                                # Razclenjevanje datuma iz stringa v timestamp
                                                ft_start_ts = datetime.datetime.strptime(split_line[0] + " " + \
                                                        split_line[1][:8], "%d/%m/%y %H:%M:%S")
                                        else:
                                                continue
        # Preverimo da ni ze konec
        if "FINISHED" in split_line:
            break
        # Konec parsanja
```

Koda 4: Razčlenjevanje dnevnika po vrsticah

```
#Zapisi v CSV le ce je prislo do spremembe v casu in je velikost bloka 750kB
                if b_start_ts != bp_start_ts_zoom and blocknr > 0 and blocknr % 15 == 0:
                    datablockzoom_array.append(datablock.copy())

                    # Zracunaj razliko v casu
                    b_diff_ts = b_start_ts - bp_start_ts_zoom

                    # Zracunajo povprecno hitrost za zdruzen block
                    avg_speed = b_size_zoom / b_diff_ts.total_seconds()
                    for tmpdata in datablockzoom_array:
                        b_elapsed_time = round(b_diff_ts.total_seconds()/ \
                                        len(datablockzoom_array), 2)
                        tmpdata['block_elapsed_time']=b_elapsed_time
                        tmpdata['block_avg_speed']=round(avg_speed, 1)
                        csvblockzoomfile.writerow(tmpdata.values())
                    b_size_zoom = 0
                    bp_start_ts_zoom = b_start_ts
                    datablockzoom_array = []
                #Konec grobega CSV-ja
```

Koda 5: Koda za izračun časa prenosa in upoštevanje 15x povečanje bloka

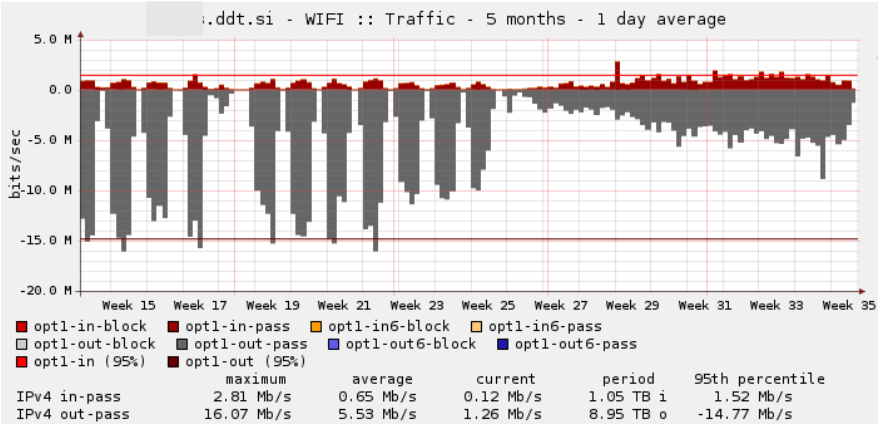
4 PREDSTAVITEV REZULTATOV

4.1 Rezultati za prvi sklop meritev

Rezultate za prvi sklop bomo predstavili s pomočjo SQL stavkov v navezi z grafi narejenimi v Microsoft Excel ter z grafi pridobljenimi iz RRD podatkov. Za prikaz rezultatov analize drugega sklopa pa smo uporabili grafe iz spletne aplikacije MicroStrategy Analytics Express.

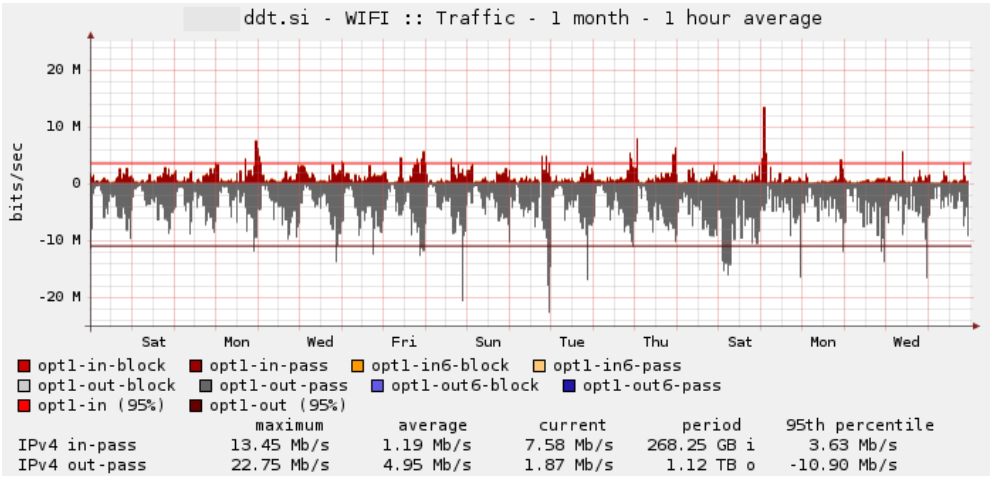
4.1.1 Analiza iz RRD grafov

Slika 20 nazorno prikazuje razliko pri uporabi brezžičnega omrežja med šolskim letom in turistično sezono. Prikazani so podatki za obdobje od 1.4.2014 do 30.8.2014. V tem obdobju je bila najvišja hitrost prenosov 16 Mb/s, prenesenih pa je bilo 8 TB podatkov.



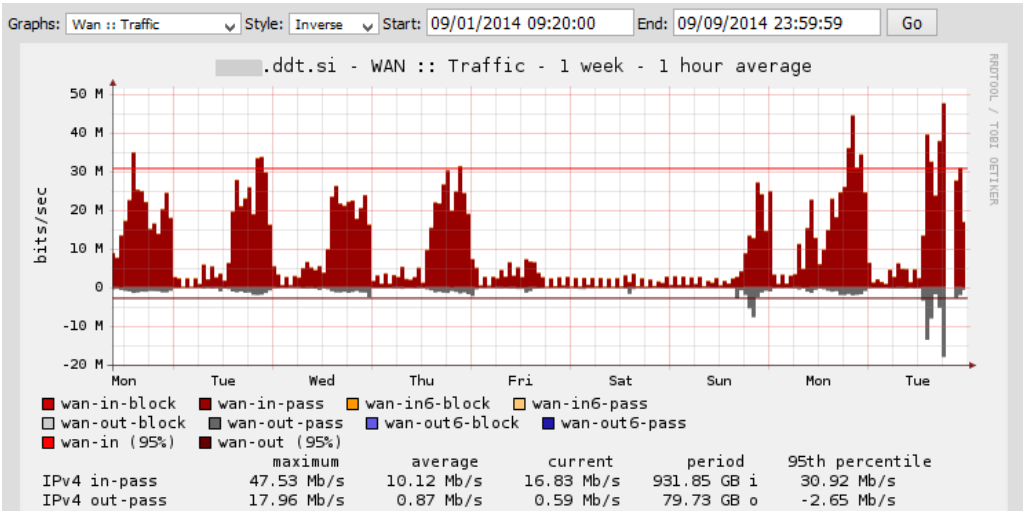
Slika 20: RRD - razlika uporabe omrežja med šolskim letom in turistično sezono

Slika 21 prikazuje obdobje, ko smo izvajali tudi naš eksperiment (8.8.2014 do 29.8.2014). V tem mesecu je bil hostel najbolj zaseden, povprečno okrog 250 gostov na dan, kar se izraža tudi v enakomerni uporabi omrežja. Povprečna hitrost prenosov je bila 5 Mb/s, gosti so skupno prenesli 1 TB podatkov, kar za takšno število uporabnikov ni ravno visoka številka.



Slika 21: RRD - prikaz povprečnih hitrosti v brezžičnem omrežju, avgust 2014

Na spodnji sliki *(slika 22)* je prikazana uporaba celotnega dijaškega omrežja v času prvega tedna šolskega leta, od 1.9.2014 do 9.9.2014. Razvidno je, da je omrežje bolj v uporabi v popoldanskem času, največja izmerjena hitrost prenosa je bila 47.5 Mb/s, povprečna hitrost je bila 10 Mb/s, kar je dovolj za običajnega uporabnika. V tem obdobju so prenesli samo 1 TB podatkov, kar je verjetno posledica blokiranja *torrent* prenosov.



Slika 22: RRD – WAN prikaz celotne uporabe omrežja v prvem tednu šolskega leta 2014/15

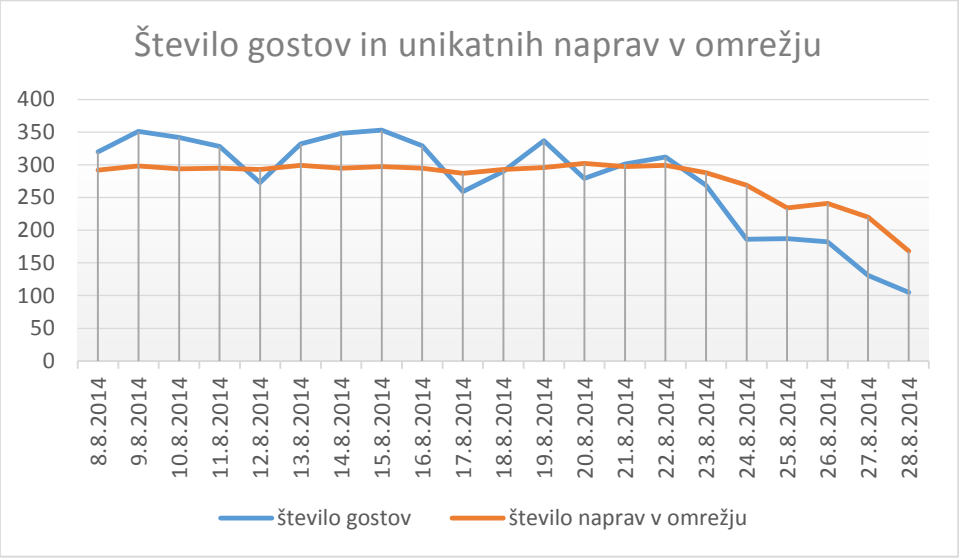
4.1.2 Analiza iz squid in požarnega dnevnika

Najprej nas je zanimalo, koliko dejansko je uporabnikov našega omrežja. S pomočjo Hostel programa za vodenje evidence gostov smo beležili dnevno število gostov, ki so pri nas prespali. Iz squid dnevnika smo z SQL stavkom (*koda 6*) pridobili podatke o unikatnih IP-jih, ki so bili aktivni na določen dan. Ko smo podatke predstavili na grafu, nas je presenetilo kako velik odstotek gostov uporablja omrežje, opazili pa smo tudi, da je na trenutke v omrežju celo več naprav kot gostov (*slika 23*). Do tega lahko pride, saj nekateri uporabljajo več kot

eno napravo (recimo telefon in tablico/prenosnik), drugo dejstvo pa je, da se je po eni noči izmenjalo preko 200 gostov. Ti gosti so lahko dostopali do omrežja v istem dnevu, kot so se odjavili iz hostla, poleg njih pa je prišlo še približno 200 novih. In tako imamo lahko takoj 400 unikatnih naprav v omrežju v enem dnevu.

```
select ts::date as date, count(distinct source_ip) as distinct_source_ip_count
from squid
group by date
order by date
```

Koda 6: SQL poizvedba za št. unikatnih naprav v omrežju

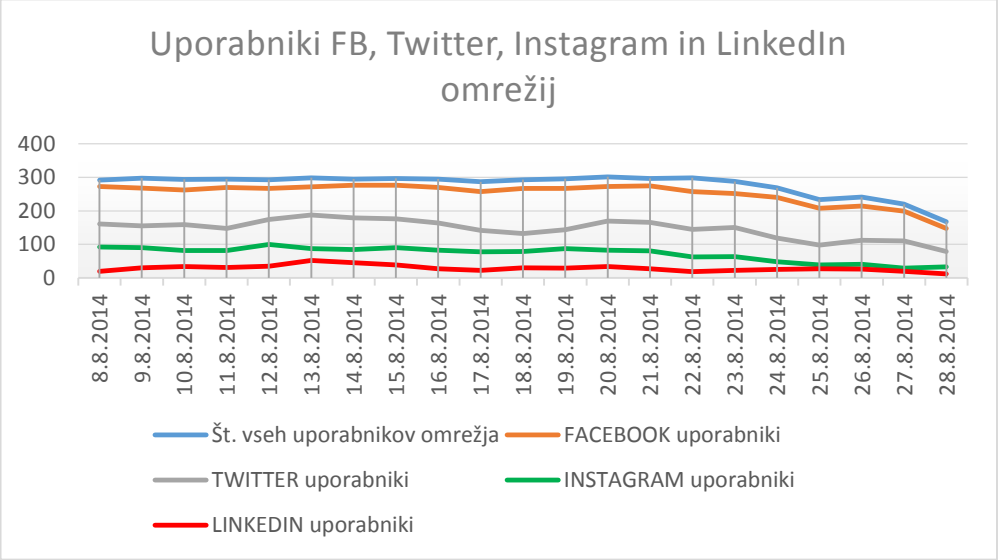


Slika 23: Število unikatnih naprav v omrežju

V naslednji obdelavi smo se posvetili uporabi popularnih omrežij, kot so Facebook, Twitter, Instagram in Linkedin. Ob dejstvu, da ima vsak pameten telefon že možnost namestitve aplikacij za dostop do teh omrežij, nas niti ne preseneča tako velika uporaba Facebook omrežja. Preostala tri so tudi manj popularna, kar se odraža tudi na grafu (*slika 24*). Analiza ni bila tako enostavna, saj se FB in Twitter dostopi vršijo preko HTTPS povezav, mi pa v squid lahko logiramo le HTTP URL-je, v požarnem zidu pa vidimo le dostope do željenih IP-jev. V ta namen smo na spletu izbrskali podatke o IP-jih, ki jih uporabljajo ta omrežja (posamezno omrežje lahko uporablja več IP-jev), celoten rang IP-jev smo zapisali v podatkovno bazo v novo tabelo *ip_range* in s pomočjo nje pridobili želene podatke. [26, 27]

```
select q.ts::date as date, count(source_ip) as request_count, count(distinct source_ip) as
distinct_source_ip_count
from ip_range ir right join qnap q on ir.range::inet >> q.destination_ip::inet
where ir.domain = 'facebook'  (isto pognali še za 'twitter', 'instagram' in 'linkedin")
group by date
order by date
```

Koda 7: SQL poizvedba za prikaz FB, TW, IG in LinkedIn omrežij

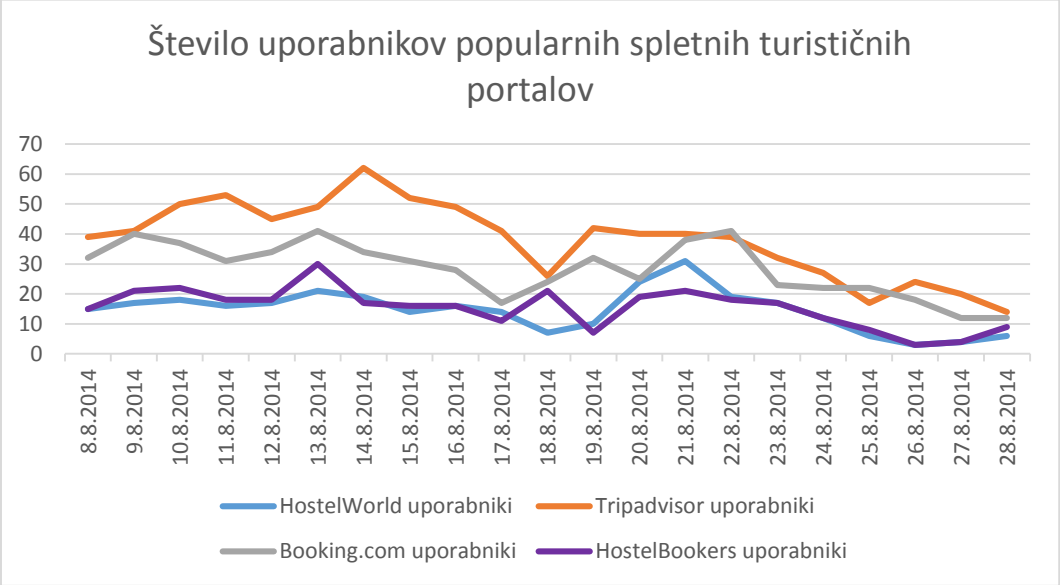


Slika 24: Število unikatnih uporabnikov FB, TW, Instagram in LinkedIn omrežij

Glede na to, da v poletnem času delujemo kot hostel, nas je zanimalo tudi v kolikšni meri uporabniki dostopajo do posameznih rezervacijskih portalov (*koda 8*). Najbolj popularna stran za pridobivanje informacij o turističnih destinacijah je očitno TripAdvisor (*slika 25*).

```
select ts::date as date, count(source_ip) as request_count, count(distinct source_ip) as
dist_source_ip_count
from squid
where domain = 'hostelbookers.com'  (isto še za: hostelworld.com, tripadvisor.com, booking.com)
group by date
order by date
```

Koda 8: SQL poizvedba za prikaz uporabnikov turističnih portalov



Slika 25: Število uporabnikov spletnih turističnih portalov

Zelo nas je zanimala tudi uporaba omrežja po urah skozi celoten mesec. Na spodnjem grafu (*slika 26*) vidimo kakšen je trend uporabe omrežja. Pogledali smo število unikatnih naprav in število zahtevkov (*angl. requests*) po spletnih vsebinah v posamezni uri (*koda 9*). Vidimo, da sta dve špici in sicer ena je sredi dneva, druga pa se začne v večernih urah in traja vse do zgodnjega jutra.

```
select extract(hour from ts) as hour_of_day, count(*)
from qnap
group by hour_of_day
order by hour_of_day

select u.hour_of_day, count(distinct u.source_ip)
from
(               select source_ip, extract(hour from ts) as hour_of_day, count(*) as request_count
                from qnap
                group by source_ip, hour_of_day
                order by hour_of_day, source_ip
) u
group by u.hour_of_day
order by u.hour_of_day
```

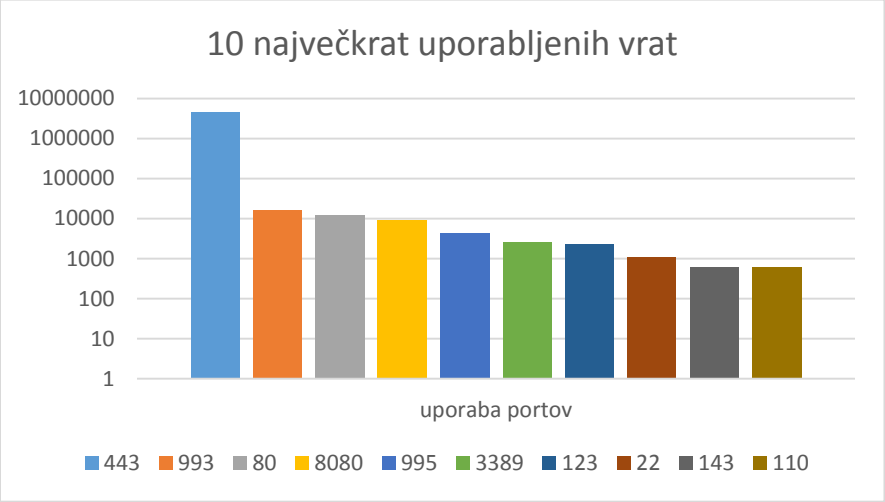
Koda 9: SQL poizvedba za št. zahtev in št. aktivnih uporabnikov po urah

Slika 26: Število zahtevkov in uporabnikov po urah

Pogledali smo si še, kateri protokoli so največ v uporabi. Z SQL poizvedbo (*koda 10*) smo prišli do podatka, da gre velika večina prometa preko zaščitenega HTTPS protokola (vrata 443 se pojavijo v dnevniku kar 4.634.321 krat), sledijo mu IMAPS (vrata 993 se pojavijo 16.555 krat), šele na tretjem mestu se pojavi HTTP promet (vrata 80 se pojavijo 12.129, vrata 8080 pa 9313 krat). Naslednji s 4.202 dostopi se pojavi POP3S protokol na vratih 995 (*slika 27*).

```
select  gateway, destination_port, count(*) AS port_count
from qnap
group by gateway, destination_port
order by port_count DESC
limit 10
```

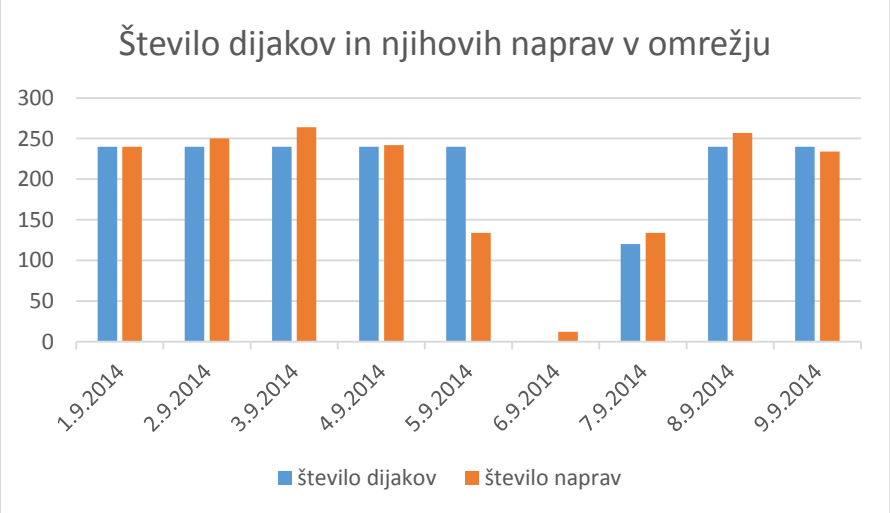
Koda 10: SQL poizvedba za prikaz 10 popularnih protokolov



Slika 27: 10 največkrat uporabljenih vrat

Pri spodnjih slikah smo analizirali uporabo omrežja in navade dijakov v prvem tednu šolskega leta (1.9.2014 – 9.9.2014). Uporabljali smo enake SQL stavke kot pri analizi omrežja med turistično sezono, le podatki so za drugo obdobje, namesto turističnih portalov pa smo si ogledali dostope do znanih slovenskih novičarskih strani.

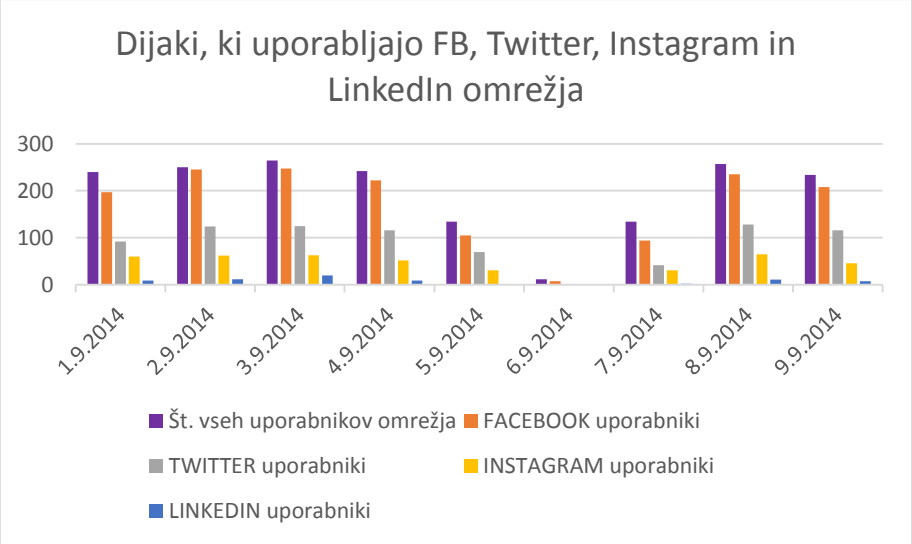
Slika 28 nam pokaže, da imamo v domu 240 dijakov, unikatnih naprav, ki so na posamezni dan v omrežju, pa je skoraj isto. To kaže na to, da imajo dijaki tudi po več naprav, v tej statistiki pa so prikazani tudi naključni gosti, ki so prišli na obisk v dijaški dom k dijakom. V nedeljo sem upošteval le ½ gostov, saj jih nekaj pride v ponedeljek.



Slika 28: Število dijakov in njihovih naprav omrežju

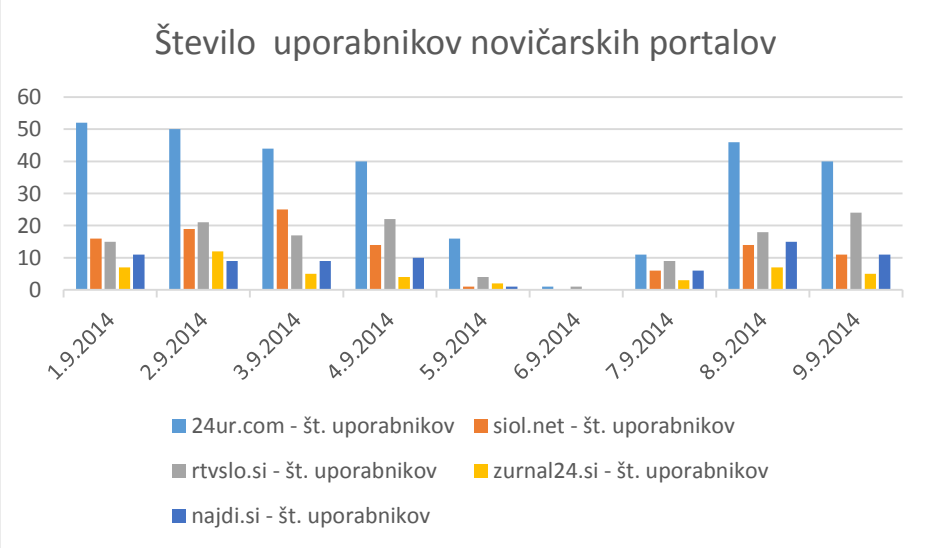
Spodaj (*slika 29*) je prikazano število vseh naprav, ki jih uporabljajo dijaki na določen dan v navezi z dostopi do socialnih omrežij (Facebook, Twitter, Instagram in LinkedIn). Po

pričakovanjih je največ dostopov do Facebook-a, dosti jih uporablja tudi Twitter in Instagram, pričakovano pa je zelo malo uporabnikov LinkedIn omrežja, saj je namenjeno bolj poslovnim uporabnikom.



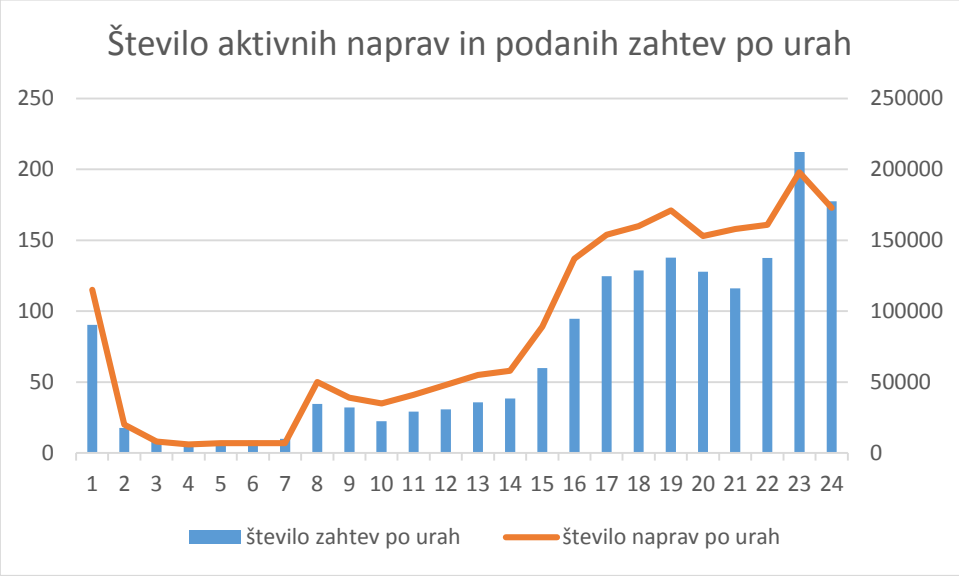
Slika 29: Število dijaških naprav, ki dostopa do FB, TW, Instagram in LinkedIn omrežij

Pri gostih hostla smo si ogledali dostope do turističnih portalov, pri dijakih pa si bomo ogledali uporabo popularnih novičarskih strani. Kot smo pričakovali, je v vodstvu portal 24ur.com, sledi mu siol.net in rtvslo.si (*slika 30*). Pričakovano je tudi majhna uporaba iskalnika najdi.si, saj je v večini brskalnikov privzet iskalnik google.si. V analizo ga nismo mogli vključit, ker se dostopi vršijo preko HTTPS povezave.



Slika 30: Dijaki - število dostopov do novičarskih portalov

Na spodnji sliki (*slika 31*) vidimo čisto drugačne navade uporabnikov v primerjavi z poletnimi gosti. Pri dijakih je lepo razvidno kdaj spijo, ter kdaj jih je večina v šoli, medtem ko je bilo omrežje v času poletja uporabljano ves čas. Večina dijakov je v šoli do okrog 13h, ko začne število zahtevkov po spletnih straneh tudi naraščati. Umirijo pa se nekje po 1h ponoči. Najvišja uporaba omrežja je pričakovano v večernem času.

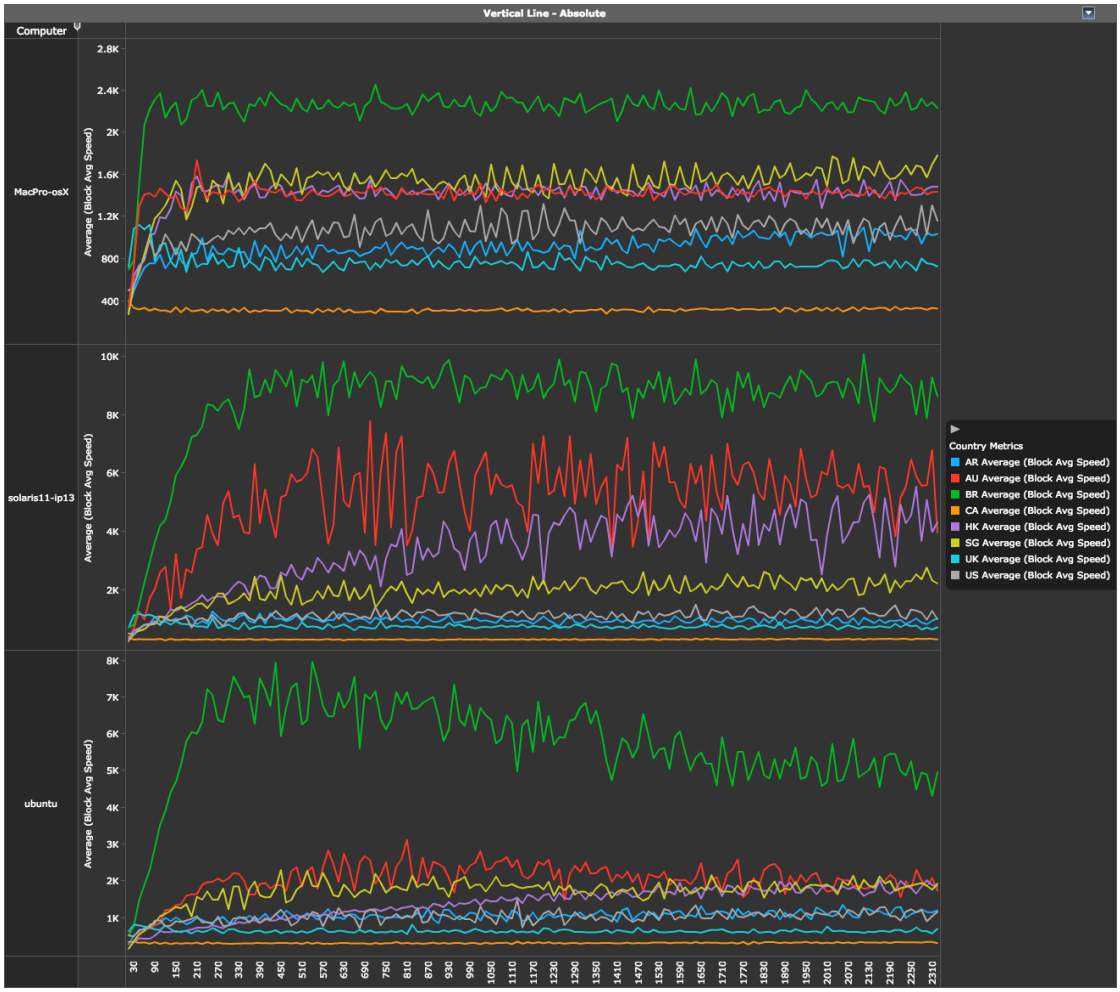


Slika 31: Dijaki - število naprav in podanih zahtev po urah

4.2 Rezultati za drugi sklop meritev

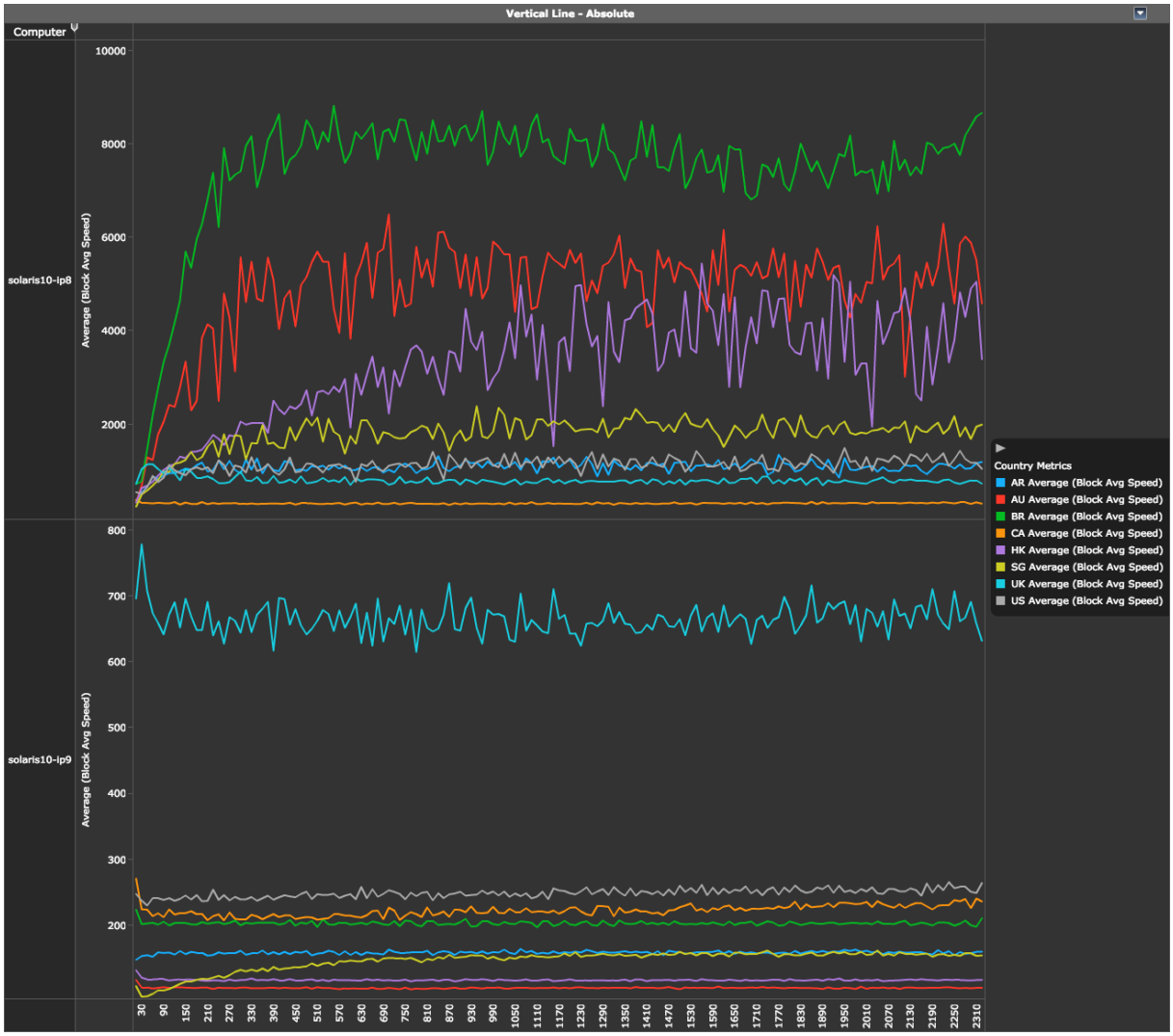
4.2.1 Povprečen prenos podatkov

Najprej si bomo pogledali analizo prenosov podatkov iz posameznih držav. Ob predstavitvi prvega grafa za prvi računalnik *Solaris11-ip13* smo bili rahlo presenečeni, saj so bili prenosi iz zelo oddaljenih držav najhitrejši. Ko pa smo enak graf pripravili še za računalnik *Solaris10-ip8*, smo ugotovili, da je najhitrejši prenos iz Velike Britanije. Z veliko zagnanostjo smo pripravili grafe še za preostale računalnike in prišli do ugotovitev, ki jih predstavlja spodnji graf (*slika 32*). Ugotovili smo, da visoka latenca nekaterih oddaljenih strežnikov omejuje prenos tako v Ubuntu kot v Mac OS X. To nakazuje na premajhen vmesni pomnilnik (*angl. buffer)* pri TCP nastavitvah. Latenca do HK in AU strežnikov je zelo podobna, vendar vidimo, da strežnik iz HK bistveno počasneje viša hitrost prenosa. Najverjetneje gre za uporabo različnega TCP *Congestion Control* algoritma (npr. *reno vs h-tcp*). [28]



Slika 32: Povprečna hitrost prenosa bloka

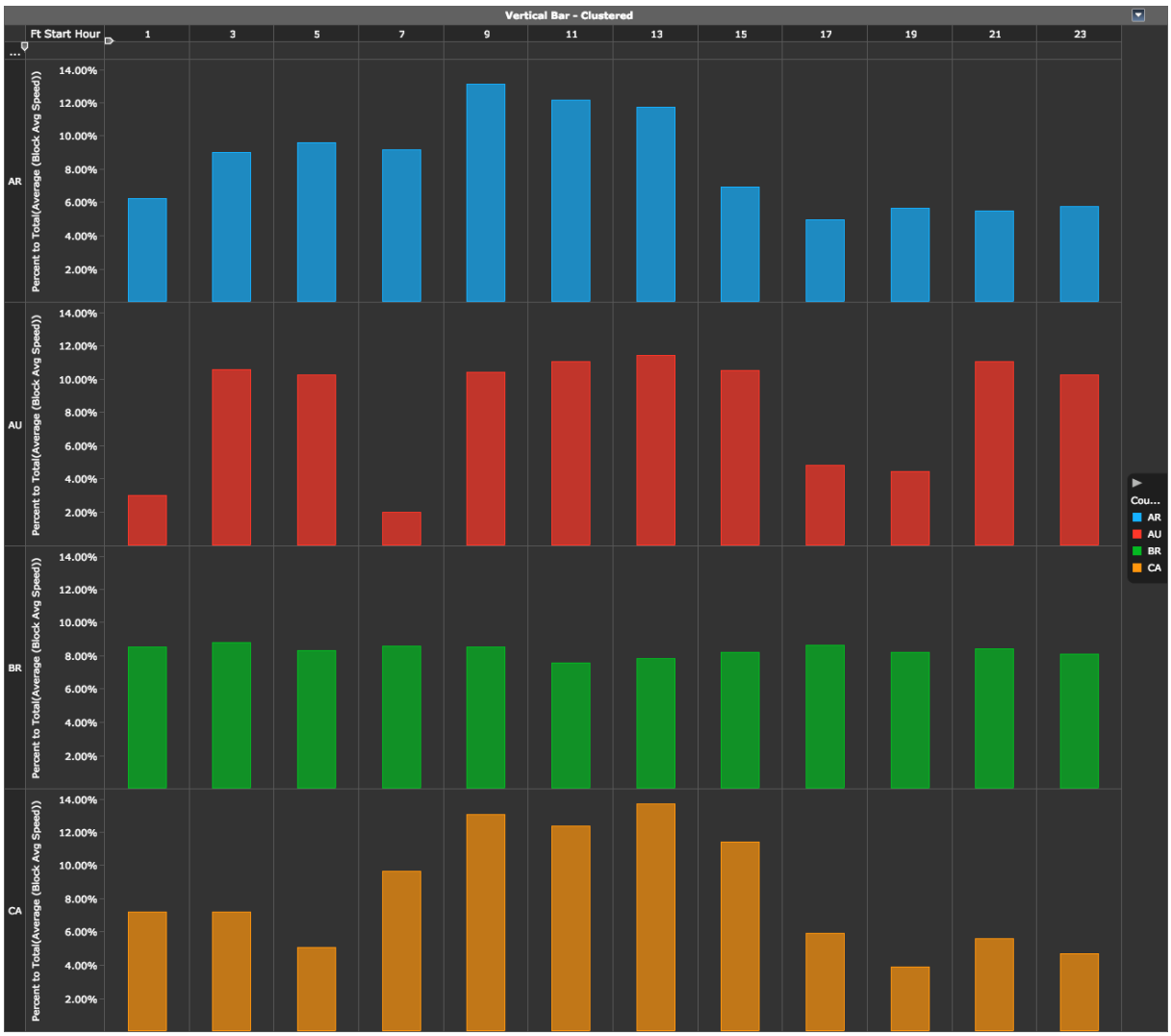
Na drugem grafu (*slika 33*) bomo videli zanimivo razliko pri različnih TCP nastavitvah na enakem operacijskem sistemu Solaris 10. Na računalniku s povečanim vmesnim pomnilnikom *(angl. buffer)* je najhitrejši prenos iz Brazilije, nato iz Hong Konga, najpočasnejši pa se je izkazal prenos iz strežnika iz Velike Britanije (prenos okrog 700 KB/s). Ko pa pogledamo prenose s privzetimi nastavitvami na računalniku *Solaris10-ip9*, pa vidimo da so prenosi iz Velike Britanije daleč najhitrejši. To dokazuje, da so privzete nastavitve na Solaris 10 neprilagojene za WAN (*angl. Wide Area Network*) omrežja, saj so TCP nastavitve *buffer-jev* bistveno premajhne in s tem je prenos omejen z latenco in ne s prepustnostjo linije.



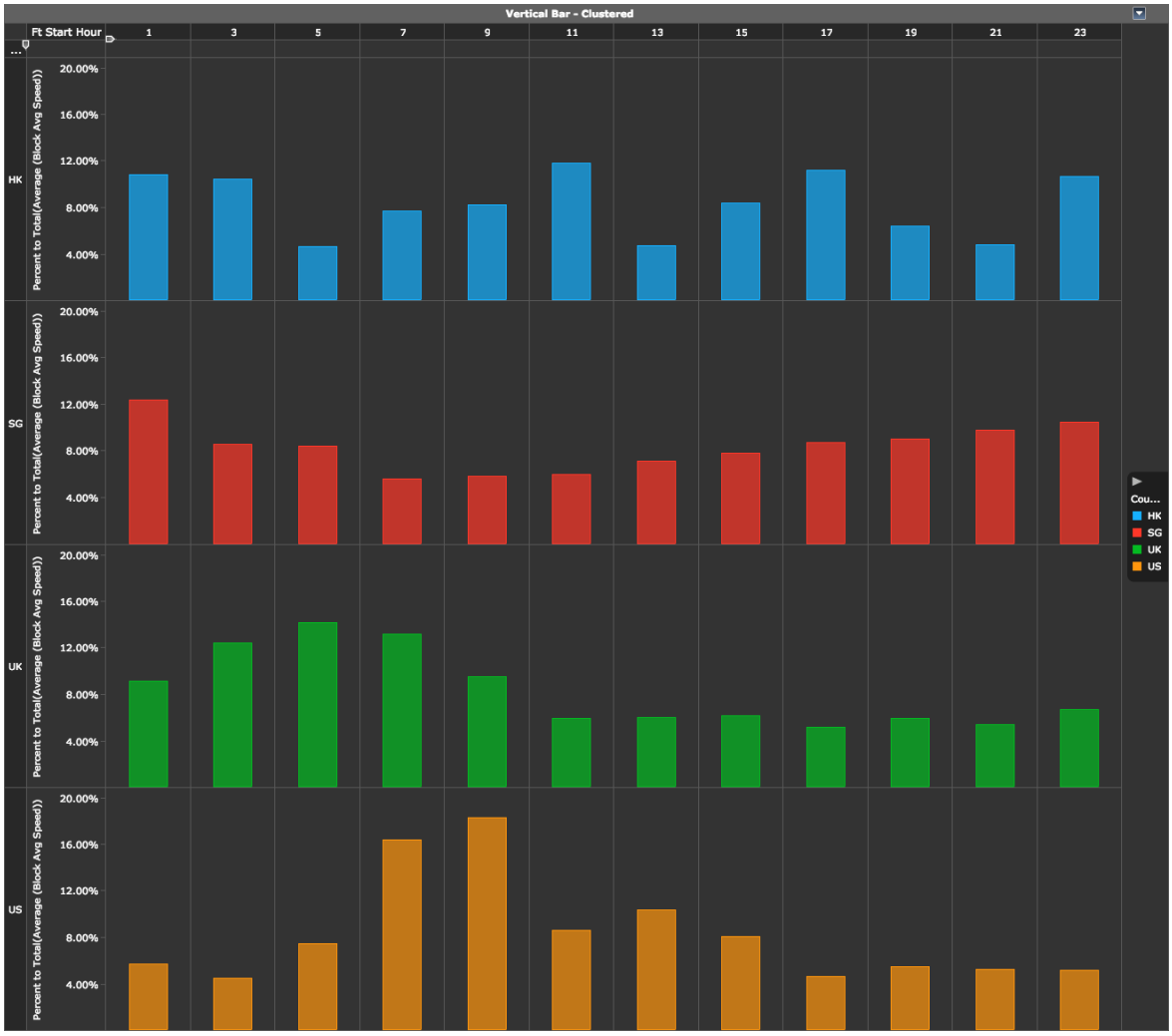
Slika 33: Povprečna hitrost prenosov na enakem OS z različnimi TCP/IP nastavitvami

4.2.2 Prenosi podatkov tekom dneva

Analiza je pokazala, da ura v dnevu nima drastičnega vpliva na hitrosti prenosov podatkov. Iz Brazilije je konsistenten čez cel dan (*slika 34*). Iz Argentine, Kanade, Velike Britanije Singapurja in ZDA je hitrost rahlo višja v nočnem času v posamezni državi (*slika 34 in slika 35*). Zanimivo je dvojno zmanjšanje hitrosti med dnevom pri prenosu iz Hong Konga (*slika 35*). Promet med Azijo in Evropo gre večinoma preko ZDA (*slika 39*), to je tudi v našem primeru pokazal program *traceroute*. Na podlagi tega ugibam, da je ena upočasnitev predstavljala zasičenost povezav na Atlantiku, druga pa na Pacifiku.



Slika 34: Prikaz prenosov po urah za AR, AU, BR in CA



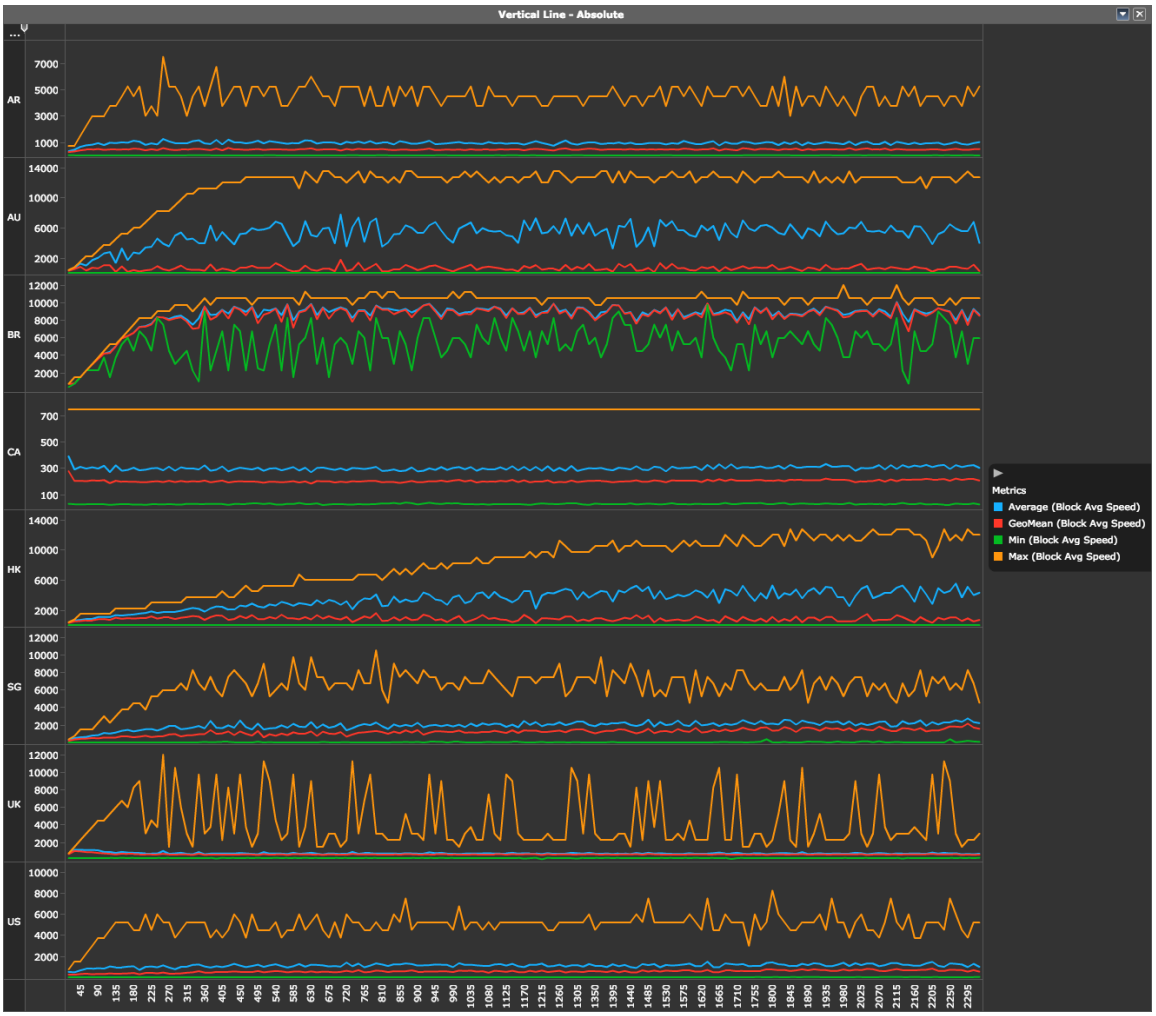
Slika 35: Prikaz prenosov po urah za HK, SG, UK in US

4.2.3 Variacija prenosov iz posameznih strežnikov

Tukaj bomo poskušali za vsak testni računalnik analizirati povprečno, najhitrejšo in najpočasnejšo hitrost prenosov podatkov ter tudi geometrijsko sredino. Pri vseh teh analizah je skupno to, da strežnik v Braziliji omogoča konsistentno hitre prenose. Glede na te enako hitre prenose je povsem možno, da ga omejuje nastavitev TCP velikost okna in vmesnega pomnilnika (*angl. buffer*) in ne sama pretočnost linije. V povprečju sta hitra tudi strežnika v Avstraliji in HK, odstopanje od maksimalnega prenosa ni veliko. Pri ostalih strežnikih pa vidimo, da je povprečna hitrost razmeroma majhna glede na maksimalno zmerjeno.

V oči nam je padla krivulja maksimalnega prenosa iz Avstralije in Hong Konga. Latenca in hitrost sta zelo podobni, vendar krivulja do dosežene maksimalne hitrosti prenosa je zelo različna. To kaže na to, da imata strežnika nastavljena dovolj velike TCP buffer-je, vendar

uporabljata različne TCP *Congestion Avoidance* algoritme. Na grafu iz Hong Konga lahko ugotovimo, da je po prenesenih 100 MB hitrost še vedno naraščala, kar pomeni, da uporabljajo bolj konzervativen algoritem. (*slika 36*)

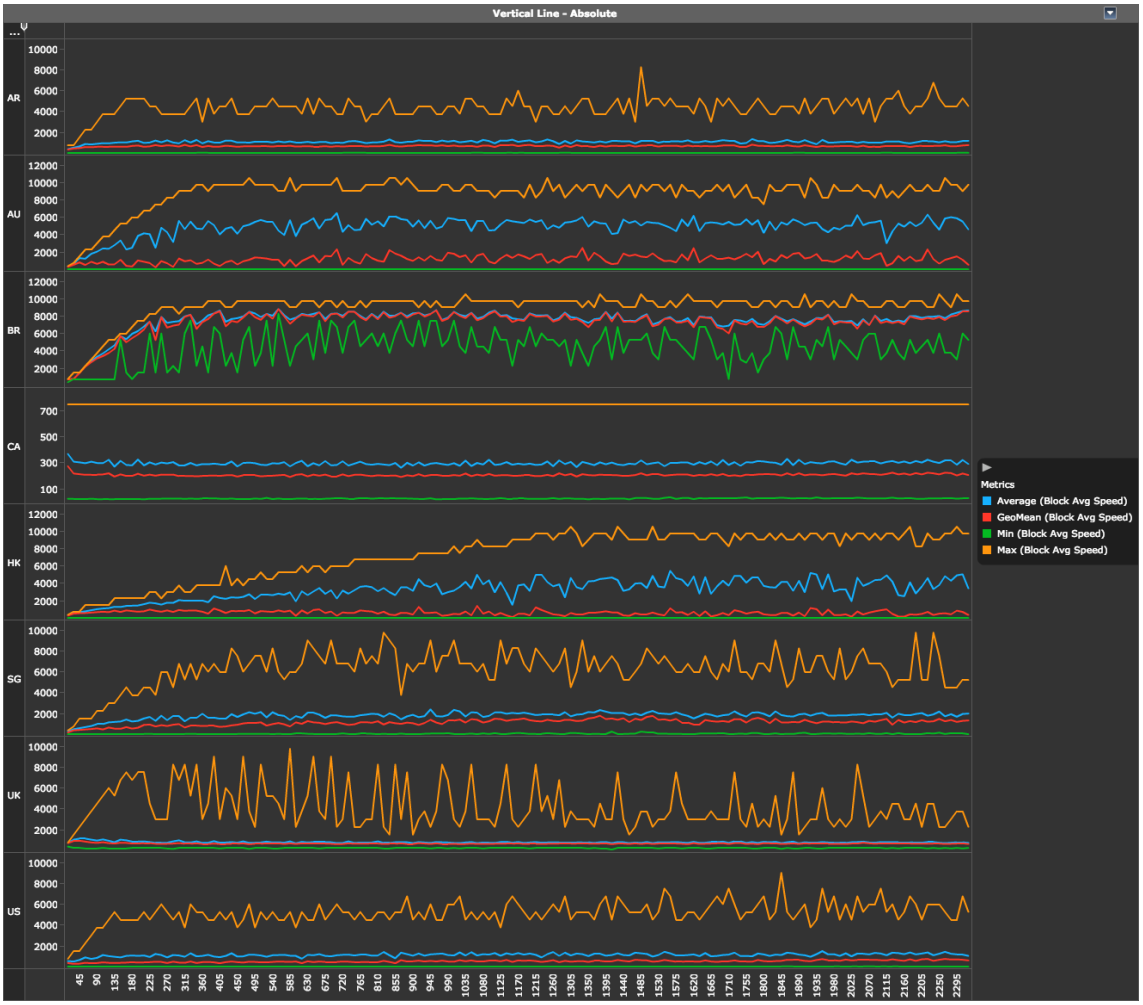


Slika 36: Solaris11-ip13 variacija prenosov

4.2.4 Vpliv latence na prenos podatkov

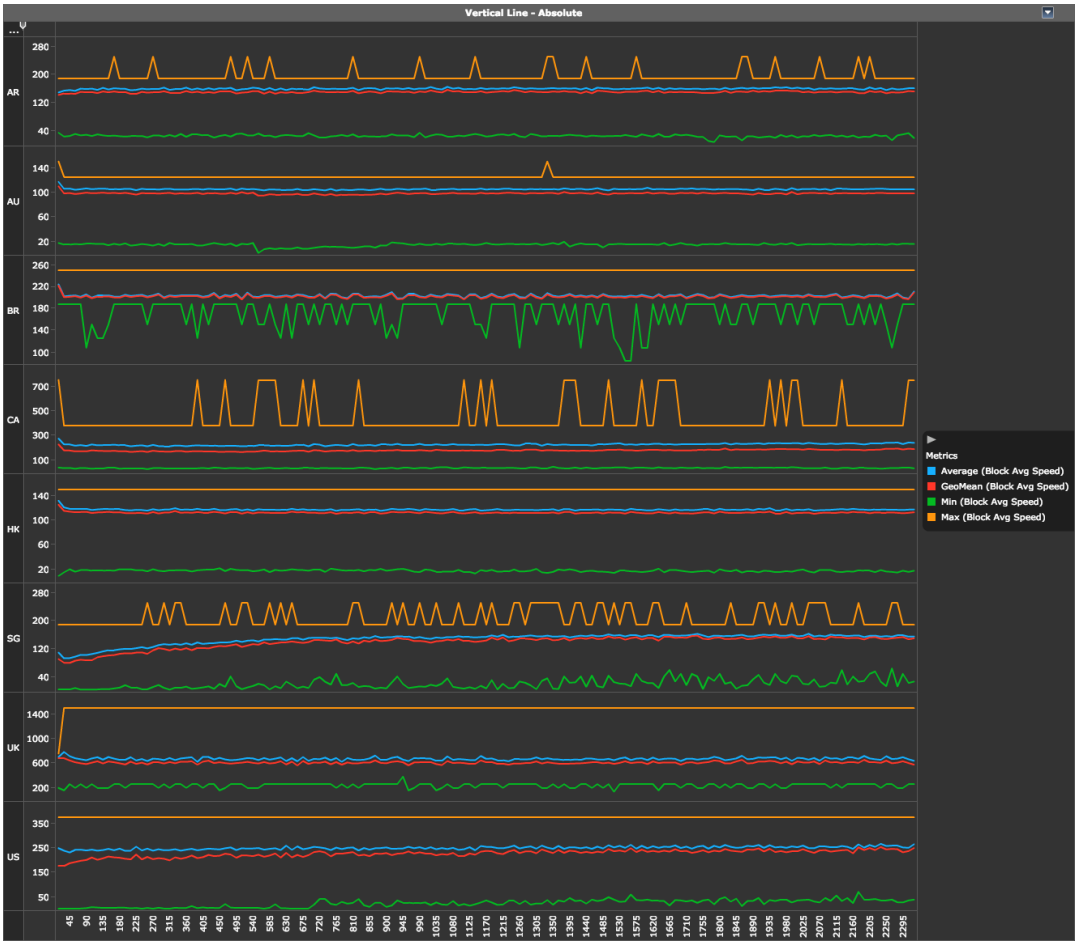
Odzivnost internetnih storitev je večkrat omejena z latenco kot pa s hitrostno omejitvijo prenosa. Negativen vpliv ima tudi na govorne in interaktivne storitve, kot je recimo oddaljen nadzor (*RDC*) in druge. Vpliv latence pri večjih datotekah lahko omejimo z ustreznim povečanjem TCP vmesnih pomnilnikov. Na spodnjih slikah je prikazan prenos datoteke na z optimiziranimi (za visoko latenco) in privzetimi TCP nastavitvami na Solaris10 sistemu.

Na Solaris10-ip8 sistemu z optimiziranimi nastavitvami (*slika 37*) se lepo vidi kako hitrost niha. Krivulje so dinamične glede na zasičenost linij, hitrosti pa so skoraj nepredstavljivo višje kot pri sistemu s privzetimi TCP nastavitvami.



Slika 37: Solaris10-ip8 variacija prenosov

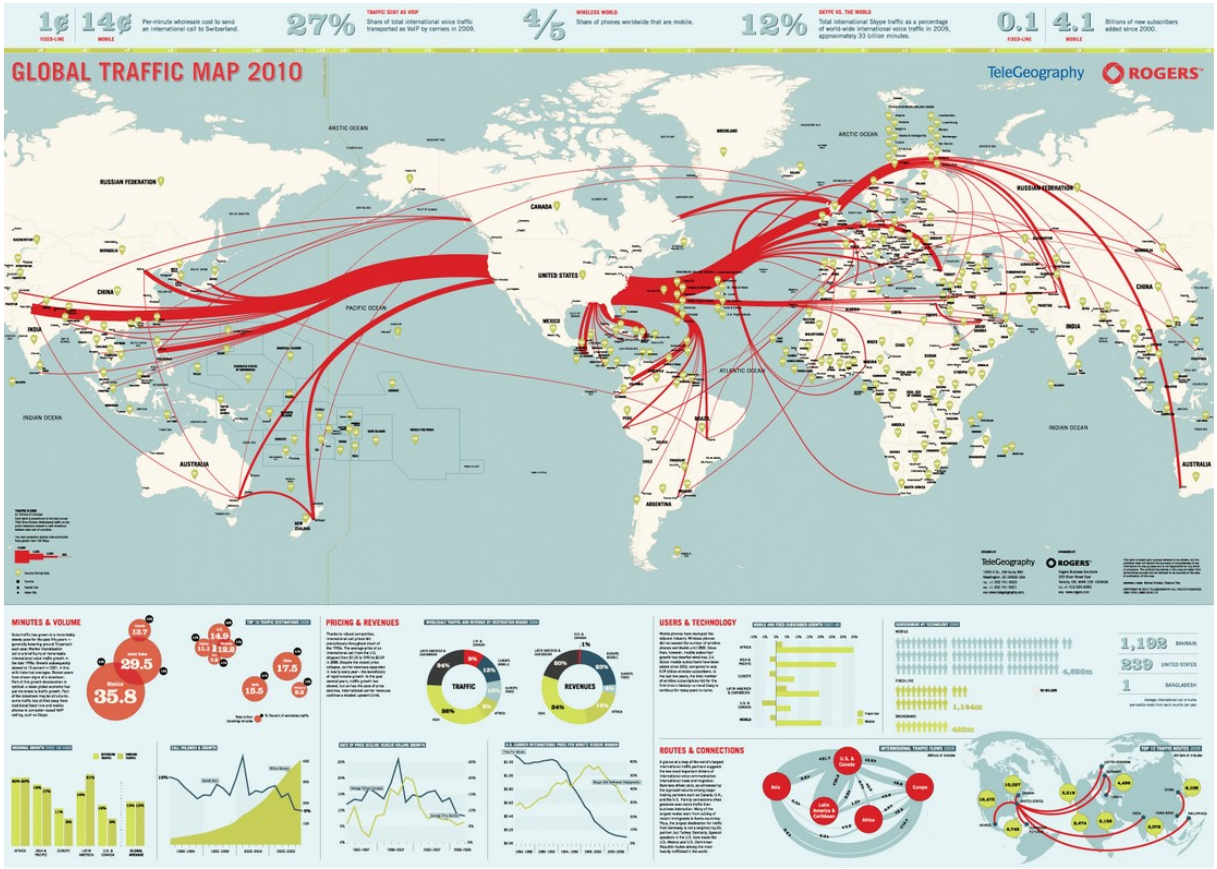
Na sistemu Solaris10-ip9 s privzetimi TCP nastavitvami so lepo vidne ravne linije pri prenosu, kar nakazuje da prenos ni omejen s prepustnostjo ampak z latenco (*slika 38*). Če bi bil omejen z latenco bi namreč hitrost nihala, saj internet ne garantira enake hitrosti prenosa, tudi navzgor omejenega ne. Tudi absolutna hitrost prenosa je nekajkrat nižja kot pri optimiziranem Solaris10-ip8 sistemu na *slika 37*. Najhitrejši so bližnji strežniki, medtem ko so tisti z visoko latenco v Avstraliji in Hong Kongu omejeni na vsega 100 KB/s.



Slika 38: Solaris10-ip9 variacija prenosov

4.2.5 Globalno omrežje

Pri iskanju strežnikov za to analizo sem pri praktično vseh Azijskih strežnikih opazil veliko latenco. *Traceroute* je pokazal, da promet do večine strežnikov v Aziji poteka preko ZDA, kar je bistveno dlje kot bi bilo po najkrajši možni poti. Po kratki raziskavi sem ugotovil, da so ZDA še vedno središče interneta, kar prikazuje spodnja slika podjetja TeleGeography (*slika 39*).



Slika 39: Prikaz globalnih povezav
(vir: http://scjsin.websandboxes.com/wp-content/uploads/2010/05/global-traffic-map-large-300x216.png)

5 SKLEPNE UGOTOVITVE

V diplomskem delu smo si ogledali omrežno infrastrukturo Dijaškega doma Tabor. Kot izobraževalna ustanova smo del javnega sektorja in kot taka precej omejeni pri financiranju infrastrukture iz lastnih sredstev. Sofinanciranje s strani države je zaradi finančne krize vse bolj omejeno, zato je tudi strategija dijaškega doma usmerjena v odprtokodne rešitve, ki z minimalnim finančnim vložkom lahko bistveno izboljšajo uporabniško izkušnjo. Takšen primer je tudi uvedba programskega usmerjevalnika pfSense v dijaškem domu.

V nalogi smo raziskali navade uporabnikov omrežja, preučili obremenjenost ter opravili analizo omrežja. Tako smo že na začetku ugotovili, da se je v času turistične sezone za šibko točko izkazalo obstoječe brezžično omrežje. Prejeli smo pritožbe gostov, da je pokritost brezžičnega omrežja v nekaterih sobah pomanjkljiva. Delno smo te težave odpravili z dodatnimi dostopnimi točkami ter porazdelitvijo obstoječih. Rezultati meritev obremenjenosti omrežja kažejo na relativno majhne prenesene količine podatkov, kar lahko delno razložimo s strukturo gostov in njihovimi potrebami. Gosti so večinoma posegali po raznih turističnih informacijah preko spleta oz. dostopali do socialnih omrežij, npr. Facebook-a. Podatkovna obremenitev tovrstnih dostopov praviloma ni velika, torej hitrosti prenosa nekaj 10 Mb/s ne presenečajo in so v skladu s pričakovanji. V času prvega tedna šolskega leta, se je prenos nekoliko povečal, kar pripisujemo večji prisotnosti dijakov, predvsem v popoldanskem času. Kljub temu povprečne hitrosti niso bistveno odstopale od turistične sezone. Deloma gre iskati vzroke v blokadi določenih protokolov, predvsem P2P omrežij. Drugi možen vzrok je premajhna prepustnost obstoječega brezžičnega omrežja, saj v nekaterih stavbah dijaškega doma uporabljamo še stare dostopovne točke s hitrostjo 54 Mb/s. Prav tako pokritost dostopnih točk ni optimalna, kar nameravamo v bližnji prihodnosti z dodatnim financiranjem urediti. Pokazal se je tudi trend, da posamezen uporabnik omrežja premore več naprav, predvsem pametne telefone ter tablico in/ali prenosni računalnik. Ta podatek je lepo razviden iz podatkov o prijavljenem številu gostov oz. prisotnih dijakov ter dejanskim številom dodeljenih unikatnih IP naslovov. Pri analiziranju navad uporabnikov in njihovih dostopov do socialnih omrežij, smo ugotovili, da slednja sovpada z dejansko priljubljenostjo in razširjenostjo posameznega omrežja. Tako ni presenečenje, da je socialno omrežje Facebook najbolj priljubljeno, sledijo pa mu Twitter, Instagram in Linkedin. Med novičarskimi stranmi do katerih dostopajo dijaki, je najbolj priljubljen portal 24ur.com, sledita mu siol.net in rtvslo.si.

Rezultati meritev medcelinskih prenosov so ovrgli naša prepričanja, da so prenosi iz oddaljenih krajev počasni zgolj zaradi nezmogljivih povezav. Z analizo smo pokazali, da niso vedno ozko grlo povezave, ampak tudi latence, ki imajo negativen vpliv tudi na druge storitve, ne samo na prenos datotek. Z ustreznimi nastavitvami TCP protokola lahko pridemo do tega, da so oddaljeni strežniki celo najhitrejši.

5.1 Možne izboljšave

Iz opravljene analize omrežja je razvidno, da rezultati sovpadajo s svetovnimi trendi, ki kažejo na vsakoletno povečevanje zahtev po prenosu podatkov. Kljub temu, da so bile meritve v tem diplomskem delu opravljene samo za kratko obdobje, lahko zaključimo, da se število uporabnikov oz. naprav, ki so priključene v splet iz leta v leto povečuje. Prav tako so internetne vsebine čedalje bolj vsebinsko bogate oz. zahtevajo večjo količino prenesenih podatkov. To je tudi eden izmed razlogov, da je potrebno v prihodnje posvetiti več pozornosti nadgradnji omrežne infrastrukture oz. povečati prepustnost in dostopnost brezžičnega omrežja, ki je glavna dostopna točka uporabnikov dijaškega doma. S tem ukrepom bomo omogočili tudi zagotavljanje dodatnih vsebin, npr. P2P omrežij, kar do sedaj zaradi obremenjenosti ni bilo mogoče. Analiza omrežja nam v tem pogledu odpira nove možnosti, saj lahko iz navad uporabnikov predvidimo bodoče trende oz. prilagodimo delovanje omrežja uporabniku. Tukaj v bodoče vidimo velik potencial v QoS storitvah oz. storitvah zagotavljanja kakovosti dostopov. Z uvedbo programskega usmerjevalnika pfSense smo to že deloma dosegli, a potencial v prihodnje je velik. Predvsem želimo zagotoviti dodatne vsebine in protokole (P2P, multimedijske vsebine, dostop do oblačnih storitev, ipd.) v brezžičnem omrežju, ob tem pa ohraniti optimalno delovanje obstoječih storitev kot so VoIP telefonija in hiter dostop do spleta.

Izboljšave pri mednarodnih prenosih vidimo predvsem v večjem prizadevanju strokovnjakov k zmanjšanju visokih latenc v največji možni meri. Glede na to, da 100 in več megabitne povezave pri končnih uporabnikih niso več redkost, bi morali proizvajalci operacijskih sistemov prilagoditi privzete nastavitve, ali pa vsaj dodati možnost da si uporabnik sam nastavi hitrost zunanje povezave in s tem ustrezno nastavi TCP parametre. Za boljšo diagnostiko povezav, bi moralo biti dostopnih več *Looking Glass* strežnikov, kjer bi lahko vsak uporabnik iz smeri strežnika proti njemu pogledal *traceroute* in *ping* informacije.

6 VIRI

[1] Dečman Dobrnjič, Olga, Kako vodeni doživljajo vodenje, Koper, Skupnost dijaških domov Slovenije, 2002

[2] http://en.wikipedia.org/wiki/IEEE_802.1Q (dostop avgust 2014)

[3] http://en.wikipedia.org/wiki/Multi-mode_optical_fiber (dostop avgust 2014)

[4] http://www.cisco.com/web/SG/solutions/smb/velocity/Switches/3560C/Downloads/datasheet_3560c_series_switches.pdf (dostop september 2014)

[5] http://www.telsey.com/upload/data_sheet/FM1024SPoE-2xG.pdf (dostop avgust 2014)

[6] http://en.wikipedia.org/wiki/OSI_model (dostop avgust 2014)

[7] http://en.wikipedia.org/wiki/Power_over_Ethernet (dostop avgust 2014)

[8] http://www.telsey.com/upload/data_sheet/I0158_GM1024S-Rev-0-5.pdf (dostop avgust 2014)

[9] http://www8.hp.com/h20195/v2/GetDocument.aspx?docname=c04284193 (dostop avgust 2014)

[10] http://www.lancom-systems.de/fileadmin/produkte/lc_l54/L-54g_EN.pdf (dostop september 2014)

[11] http://www.ubnt.com/downloads/datasheets/unifi/UniFi_AP_DS.pdf (dostop september 2014)

[12] http://www.vmware.com/products/vsphere-hypervisor/ (dostop avgust 2014)

[13] http://en.wikipedia.org/wiki/PfSense (dostop avgust 2014)

[14] https://www.pfsense.org (dostop september 2014)

[15] Matt Williamson, pfSense 2 Cookbook, marec 2001, stran 201

[16] http://pfsensesetup.com/packet-filter-the-engine-of-pfsense (dostop september 2014)

[17] http://en.wikipedia.org/wiki/Syslog (dostop september 2014)

[18] https://fasterdata.es.net/host-tuning/background (dostop avgust 2014)

[19] https://fasterdata.es.net/host-tuning/other/ (dostop avgust 2014)

[20] Tcp_max_buf - http://docs.oracle.com/cd/E19120-01/open.solaris/819-2724/chapter4-45/index.html (dostop avgust 2014)

[21] Tpc_cwnd_max - http://docs.oracle.com/cd/E19120-01/open.solaris/819-2724/chapter4-46/index.html (dostop avgust 2014)

[22] Tcp_recv_hiwat - http://docs.oracle.com/cd/E19120-01/open.solaris/819-2724/chapter4-43/index.html (dostop avgust 2014)

[23] Tcp_xmit_hiwat - http://docs.oracle.com/cd/E19120-01/open.solaris/819-2724/chapter4-42/index.html (dostop avgust 2014)

[24] http://en.wikipedia.org/wiki/Wget (dostop avgust 2014)

[25] http://www.microstrategy.com/us/ (dostop avgust 2014)

[26] https://developers.facebook.com/docs/sharing/best-practices#crawl (dostop september 2014)

[27] Twitter IP-ji: http://bgp.he.net/AS13414#_prefixes (dostop september 2014)

[28] http://en.wikipedia.org/wiki/TCP_congestion-avoidance_algorithm (dostop september 2014)